Adamou et al
cited with Record

Pneumococcal histidine triad protein E precursor 1039
(Hypothetical AA
protein SP1004) [phtE] [Streptococcus pneumoniae] align

Score = 2016 bits (5224), Expect = 0.0
Identities = 1004/1039 (96%), Positives = 1004/1039 (96%)

```
Query: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60
            MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS
Sbjct: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60

Query: 61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120
            QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN
Sbjct: 61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120

Query: 121  EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180
            EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY
Sbjct: 121  EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180

Query: 181  TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQLS 240
            TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIP                       NMQPSQLS
Sbjct: 181  TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLS 240

Query: 241  YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300
            YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
Sbjct: 241  YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300

Query: 301  TPNGVAIPHGDHYHFIPYSKLSALEEKIARMVPISGTGSTVSTNAKPNEVVXXXXXXXXX 360
            TPNGVAIPHGDHYHFIPYSKLSALEEKIARMVPISGTGSTVSTNAKPNEVV
Sbjct: 301  TPNGVAIPHGDHYHFIPYSKLSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSN 360

Query: 361  XXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420
                   KELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA
Sbjct: 361  PSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420

Query: 421  TPSPSLPINPGTSHEKHEEDGYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKA 480
            TPSPSLPINPGTSHEKHEEDGYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKA
Sbjct: 421  TPSPSLPINPGTSHEKHEEDGYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKA 480

Query: 481  AQKHLEEVKTSHNGLDSLSSHEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540
            AQKHLEEVKTSHNGLDSLSSHEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN
Sbjct: 481  AQKHLEEVKTSHNGLDSLSSHEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540

Query: 541  KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600
            KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
Sbjct: 541  KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600

Query: 601  VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660
            VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG
Sbjct: 601  VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660

Query: 661  EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720
            EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA
Sbjct: 661  EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720

Query: 721  GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780
            GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK
Sbjct: 721  GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780
```

```
Query:  781  IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENLKLDEKVEEPKTS  840
             IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENLKLDEKVEEPKTS
Sbjct:  781  IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENLKLDEKVEEPKTS  840

Query:  841  EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP  900
             EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
Sbjct:  841  EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP  900

Query:  901  SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP  960
             SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP
Sbjct:  901  SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP  960

Query:  961  ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI  1020
             ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI
Sbjct:  961  ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI  1020

Query:  1021 ELRLPSGEVIKKNLSDLIA  1039
             ELRLPSGEVIKKNLSDLIA
Sbjct:  1021 ELRLPSGEVIKKNLSDLIA  1039
```

tr Q8DQ07   **Pneumococcal histidine triad protein E [phtE]**   1039
   Q8DQ07_STRR6 **[Streptococcus**   AA
                 pneumoniae (strain ATCC BAA-255 / R6)]   align

```
 Score = 2006 bits (5196), Expect = 0.0
 Identities = 998/1039 (96%), Positives = 1000/1039 (96%)

Query:  1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS  60
             MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS
Sbjct:  1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS  60

Query:  61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN  120
             QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN
Sbjct:  61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN  120

Query:  121  EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY  180
             EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY
Sbjct:  121  EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY  180

Query:  181  TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLS  240
             TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIP                    NMQPSQLS
Sbjct:  181  TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLS  240

Query:  241  YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR  300
             YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
Sbjct:  241  YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR  300

Query:  301  TPNGVAIPHGDHYHFIPYSKLSALEEKIARMVPISGTGSTVSTNAKPNEVVXXXXXXXXX  360
             TPNGVAIPHGDHYHFIPYSKLSALEEKIAR VPISGTGSTVSTNAKPNEVV
Sbjct:  301  TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVSSLGSLSSN  360

Query:  361  XXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA  420
                    KELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA
Sbjct:  361  PSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA  420
```

```
Query:  421  TPSPSLPINPGTSHEKHEEDGYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKA  480
             TPSPSLPINPG SHEKHEEDGYGFDANRIIAEDESGF+MSHG+HNHYFFKKDLTEEQIKA
Sbjct:  421  TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA  480

Query:  481  AQKHLEEVKTSHNGLDSLSSHEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN  540
             AQKHLEEVKTSHNGLDSLSSHEQDYP NAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN
Sbjct:  481  AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN  540

Query:  541  KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV  600
             KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
Sbjct:  541  KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV  600

Query:  601  VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG  660
             VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG
Sbjct:  601  VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG  660

Query:  661  EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA  720
             EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA
Sbjct:  661  EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA  720

Query:  721  GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK  780
             GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK
Sbjct:  721  GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK  780

Query:  781  IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENLKLDEKVEEPKTS  840
             IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQEN KLDEKVEEPKTS
Sbjct:  781  IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS  840

Query:  841  EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP  900
             EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
Sbjct:  841  EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP  900

Query:  901  SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP  960
             SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP
Sbjct:  901  SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP  960

Query:  961  ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020
             ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI
Sbjct:  961  ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020

Query: 1021  ELRLPSGEVIKKNLSDLIA 1039
             ELRLPSGEVIKKNLSDLIA
Sbjct: 1021  ELRLPSGEVIKKNLSDLIA 1039
```

tr Q6WNQ7   **Surface protein BVH-3 [bvh-3] [Streptococcus**   1039
  Q6WNQ7_STRPN   **pneumoniae]**   AA
                                                                          align

```
 Score = 2006 bits (5196), Expect = 0.0
 Identities = 998/1039 (96%), Positives = 1000/1039 (96%)

Query:    1  MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS   60
             MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS
Sbjct:    1  MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS   60
```

```
Query:  61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN  120
             QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN
Sbjct:  61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN  120

Query: 121   EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY  180
             EVKGGYIIKVDGKYYVYLKDAAHADNVPTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY
Sbjct: 121   EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY  180

Query: 181   TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXNMQPSQLS  240
             TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIP                    NMQPSQLS
Sbjct: 181   TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLS  240

Query: 241   YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR  300
             YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
Sbjct: 241   YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR  300

Query: 301   TPNGVAIPHGDHYHFIPYSKLSALEEKIARMVPISGTGSTVSTNAKPNEVVXXXXXXXXX  360
             TPNGVAIPHGDHYHFIPYSKLSALEEKIAR VPISGTGSTVSTNAKPNEVV
Sbjct: 301   TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVSSLGSLSSN  360

Query: 361   XXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA  420
                    KELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA
Sbjct: 361   PSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA  420

Query: 421   TPSPSLPINPGTSHEKHEEDGYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKA  480
             TPSPSLPINPG SHEKHEEDGYGFDANRIIAEDESGF+MSHG+HNHYFFKKDLTEEQIKA
Sbjct: 421   TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA  480

Query: 481   AQKHLEEVKTSHNGLDSLSSHEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN  540
             AQKHLEEVKTSHNGLDSLSSHEQDYP NAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN
Sbjct: 481   AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN  540

Query: 541   KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV  600
             KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
Sbjct: 541   KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV  600

Query: 601   VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG  660
             VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG
Sbjct: 601   VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG  660

Query: 661   EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA  720
             EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA
Sbjct: 661   EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA  720

Query: 721   GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK  780
             GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK
Sbjct: 721   GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK  780

Query: 781   IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENLKLDEKVEEPKTS  840
             IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQEN KLDEKVEEPKTS
Sbjct: 781   IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS  840

Query: 841   EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP  900
             EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
Sbjct: 841   EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP  900

Query: 901   SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP  960
```

```
                 SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP
Sbjct:  901      SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960

Query:  961      ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020
                 ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI
Sbjct:  961      ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020

Query:  1021     ELRLPSGEVIKKNLSDLIA 1039
                 ELRLPSGEVIKKNLSDLIA
Sbjct:  1021     ELRLPSGEVIKKNLSDLIA 1039
```

```
tr  Q6WNQ5          Surface protein BVH-3 (Fragment) [bvh-3] [Streptococcus   1019
    Q6WNQ5_STRPN     pneumoniae]                                              AA
                                                                              align
```

```
 Score = 1974 bits (5115), Expect = 0.0
 Identities = 981/1019 (96%), Positives = 981/1019 (96%)

Query:  21       CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV 80
                 CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV
Sbjct:  1        CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV 60

Query:  81       TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD 140
                 TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD
Sbjct:  61       TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD 120

Query:  141      AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA 200
                 AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA
Sbjct:  121      AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA 180

Query:  201      YIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSK 260
                 YIVPH GHYHYIP                      NMQPSQLSYSSTASDNNTQSVAKGSTSK
Sbjct:  181      YIVPHRGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSSTASDNNTQSVAKGSTSK 240

Query:  261      PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK 320
                 PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK
Sbjct:  241      PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK 300

Query:  321      LSALEEKIARMVPISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXXXKELSSASDGYIFN 380
                 LSALEEKIARMVPISGTGSTVSTNAKPNEVV                    KELSSASDGYIFN
Sbjct:  301      LSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSNPSSLTTSKELSSASDGYIFN 360

Query:  381      PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGTSHEKHEED 440
                 PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGTSHEKHEED
Sbjct:  361      PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGTSHEKHEED 420

Query:  441      GYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS 500
                 GYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS
Sbjct:  421      GYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS 480

Query:  501      HEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID 560
                 HEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID
Sbjct:  481      HEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID 540

Query:  561      EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ 620
```

```
              EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ
Sbjct: 541    EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ 600

Query: 621    KRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT 680
              KRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT
Sbjct: 601    KRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT 660

Query: 681    FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL 740
              FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL
Sbjct: 661    FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL 720

Query: 741    VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE 800
              VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE
Sbjct: 721    VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE 780

Query: 801    VPILEKENQTDKPSILPQFKRNKAQENLKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTL 860
              VPILEKENQTDKPSILPQFKRNKAQEN K DEKVEEPKTSEKVEKEKLSETGNSTSNSTL
Sbjct: 781    VPILEKENQTDKPSILPQFKRNKAQENSKFDEKVEEPKTSEKVEKEKLSETGNSTSNSTL 840

Query: 861    EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN 920
              EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN
Sbjct: 841    EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN 900

Query: 921    GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM 980
              GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM
Sbjct: 901    GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM 960

Query: 981    LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA 1039
              LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA
Sbjct: 961    LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA 1019
```

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file sp103.res made by sdavid on Wed 28 Nov 101 11:51:05-PST.

Query sequence being compared:

Number of sequences searched: SP103 (1-447)
Number of scores above cutoff: 1
                                                      1

Results of the initial comparison of SP103 (1-447) with:
File : US09471255-2.pep


```
S           0    5-   10-        50-          -100-
C
O
R
E
 0 -|
 F  |
 O  |
 R  |
 E  |
 B  |
 M  |
 U  |
 N  |
    |
SCORE  0    50    99    149    199    248    298    348    397    447
STDEV  0 ---|-----|-----|-----|-----|-----|-----|-----|-----|-----|
```

> O <   Intelligenetics
O | O
> O <

Similarity matrix   PAM-150
Threshold level of sim.   16%
                         K-tuple   2
                      PARAMETERS

match penalty   1
                            1   Joining penalty   20
Gap penalty   1.00   Window size   32
Gap size penalty   0.05
Cutoff score   0
Randomization group   0

                    SEARCH STATISTICS

Scores: Mean   Median   Standard Deviation
                 447   0   0.00

Times: CPU   Total Elapsed
               00:00:00.00   00:00:00.00

Number of residues: 1039
Number of sequences searched: 1
Number of scores above cutoff: 1

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Init. Opt.

Sequence Name   Description   Length Score Score Sig. Frame
--------------- -----------   ------ ----- ----- ---- -----
1. US09471255-2 sequence 2, application us/09   1039   447   447   0.00   0

1. SP103 (1-447)
   US09471255-2 sequence 2, application us/09471255 application 09/471,255

sequence 2, application us/09471255
general information:
   applicant: biochem pharma inc.
   applicant: hamel, jos,e
   applicant: brodeur, bernard r.
   applicant: pineau, isabelle
   applicant: martin, denis
   applicant: rioux, clement
   title of invention: novel streptococcus antigens
   file reference: 12806-11pct
   current application number: us/09/471,255
   current filing date: 1999-12-23
   prior application number: us 60/113,800
   prior filing date: 1998-12-23
   number of seq id nos: 102
   software: fastseq for windows version 3.0
   seq id no 2.
      length: 1039
      type: prt
      organism: s. pneumoniae

   (linear sequence)
   us-09-471-255-2 Length: 1039 November 28, 2001 11:45 Type: P Check: 8890 ..

Initial Score   =   447   Optimized Score =   447   Significance =   0.00
Residue Identity =   100%   Matches   =   447*   Mismatches   =   0
Gaps   =   0   Conservative Substitutions   =   0

```
                10         20         30         40         50         60
       MKFSRKYIAAGSAVIVSLSLCAYALNQHRSQENKDNRVSYVDGSQSOSSQKSENLZPDQVSQKEGIQAEQIVI
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       MKFSRKYIAAGSAVIVSLSLCAYALNQHRSQENKDNRVSYVDGSQSOSSQKSENLZPDQVSQKEGIQAEQIVI
                10         20         30         40         50         60

50         60         70         80         90        100        110        120
KITDQGYVTSHGDHYHYYNGKVPYDALFSEEELLMKDPNYQLRDADIVNEVKGGYIIKVDGRKYYVYLKDAAHA
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
KITDQGYVTSHGDHYHYYNGKVPYDALFSEEELLMKDPNYQLRDADIVNEVKGGYIIKVDGRKYYVYLKDAAHA
    80         90        100        110        120        130        140

130        140        150        160        170        180        190
DNVRTKDEINROKQEHVKDNEKVNSNVAVARSOGRYYTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSD
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DNVRTKDEINROKQEHVKDNEKVNSNVAVARSOGRYYTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSD
  150        160        170        180        190        200        210

200        210        220        230        240        250        260
LSASBLAAAKAHLAGKRNMQPSQLSYSSTASDNNTOSVARGSTSKPANKSENLOSLLKELYDSPSAQRYSESD
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
LSASBLAAAKAHLAGKRNMQPSQLSYSSTASDNNTOSVARGSTSKPANKSENLOSLLKELYDSPSAQRYSESD
  220        230        240        250        260        270        280

270        280        290        300        310        320        330
GLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARMVPISGTGTSTVSTNAKPNEVVSSLGSLSSN
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARMVPISGTGTSTVSTNAKPNEVVSSLGSLSSN
  290        300        310        320        330        340        350        360

340        350        360        370        380        390        400
PSSLITSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQICGQPTLPNNSLATPSPSLPINPGT
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
PSSLITSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQICGQPTLPNNSLATPSPSLPINPGT
  370        380        390        400        410        420        430

410        420        430        440
                                   X
```
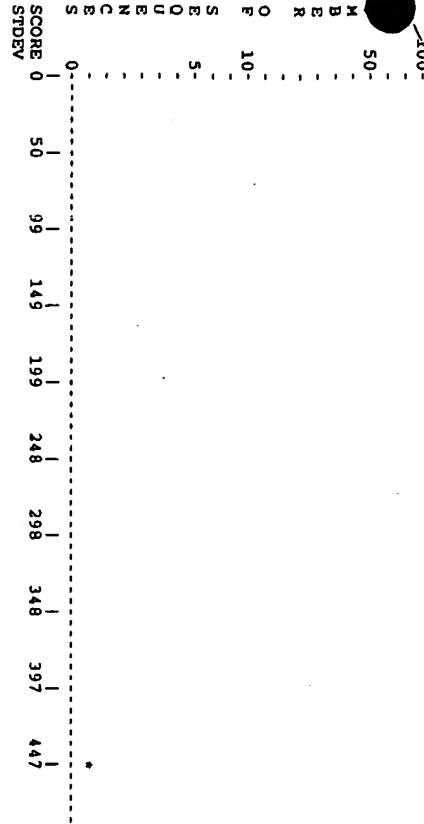
US 09/471,255

```
Query Match                 80.7%;  Score 574;  DB 4;  Length 796;
Best Local Similarity       80.6%;  Pred. No. 4.3e-56;
Matches  104;  Conservative  16;  Mismatches    9;  Indels    0;  Gaps    0;

Qy     8  AYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYVT  67
          :| |   ::::  |:|||||:|| |::|:||||||:||::||| |||||||||||||||
Db     1  SYELGLYQARTVKENNRVSYIDGKQATQKTENLTPDEVSKREGINAEQIVIKITDQGYVT  60

Qy    68  SHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKDA  127
          ||||||||||||||||||: ||||||||||||||||||||:|||||||||||||||||||
Db    61  SHGDHYHYYNGKVPYDAIISEELLMKDPNYKLKDEDIVNEVKGGYVIKVDGKYYVYLKDA  120

Qy   128  AHADNVRTK  136
          |||||||||
Db   121  AHADNVRTK  129
```

SEQ ID 56 WO98/18930

```
Query Match                 80.9%;  Score 575.5;  DB 4;  Length 763;
Best Local Similarity       79.6%;  Pred. No. 2.7e-56;
Matches  103;  Conservative  18;  Mismatches    9;  Indels    1;  Gaps    1;

Qy     7  CAYALNQHRS-QENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGY  65
          |:| |  :|:: |  |::|||||:|| |: ||:||||||:||::||| ||| |||||||||
Db     1  CSYELGRHQAGQVKKESNRVSYIDGDQAGQKAENLTPDEVSKREGINAEQXVIKITDQGY  60

Qy    66  VTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLK  125
          |||||||||||||||||||: ||||||||||||||||||:|||||:||||:|||:|||||||
Db    61  VTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYVYLK  120

Qy   126  DAAHADNVRTK  136
          |||||||:|||
Db   121  DAAHADNIRTK  131
```

WO98/18930 SEQ ID 66

US006582706B1

(12) United States Patent
Johnson et al.

(10) **Patent No.: US 6,582,706 B1**
(45) **Date of Patent: Jun. 24, 2003**

(54) **VACCINE COMPOSITIONS COMPRISING STREPTOCOCCUS PNEUMONIAE POLYPEPTIDES HAVING SELECTED STRUCTURAL MOTIFS**

(75) Inventors: **Leslie S. Johnson**, Germantown, MD (US); **John E. Adamou**, Rockville, MD (US)

(73) Assignee: **MedImmune, Inc.**, Gaithersburg, MD (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/468,656**

(22) Filed: **Dec. 21, 1999**

Related U.S. Application Data

(60) Provisional application No. 60/113,048, filed on Dec. 21, 1998.

(51) Int. Cl.^{7} A61K 39/09
(52) U.S. Cl. 424/244.1; 424/184.1; 424/185.1; 424/190.1; 424/237.1; 435/69.1; 435/320.1; 530/350; 536/23.1; 536/23.7
(58) Field of Search 424/184.1, 185.1, 424/190.1, 237.1, 244.1; 435/69.1, 320.1; 514/94; 530/350; 536/23.1, 23.7

(56) **References Cited**

U.S. PATENT DOCUMENTS

6,042,838 A * 3/2000 Briles et al. 424/244.1

FOREIGN PATENT DOCUMENTS

WO	WO 95/06732	3/1995
WO	WO 97/48417	6/1996
WO	WO 97/41151	11/1997
WO	WO-98/18930 A2 *	5/1998
WO	WO 98/18930	5/1998
WO	WO 98/18931	5/1998
WO	WO 99/15675	4/1999
WO	WO 00/17370	3/2000
WO	WO 00/39299	7/2000

OTHER PUBLICATIONS

Paul et al. Fundamental Immunology, Raven Press, New York, NY (1993) 3rd Edition, p. 251.*

Riffkin et al. A single amino–acid change between the antigenically different extracellular serine proteases V2 and B2 from Dichelobacter nodosus. Gene (1995) vol. 167, pp. 279–283.*

Abaza et al. Effects of amino acid substitutions outside an antigenic site on protein binding to monoclonal antibodies of predetermined specificity obtained by peptide immunization. Journal of Protein Chemistry (1992) vol. 11, No. 5, pp. 433–444.*

Paul W.E. In Fundamental Immunology (1993) Raves Press, New York, pp. 249–251.*

Ristori et al. Compositional bias and mimicry toward the non–self proteome in immunodominant T cell epeitpopes of self and nonself antigens. FASEB Journal (2000) vol. 14, pp. 431–438.*

Cundell, et al., "Receptor specificity of adherence of Streptococcus pneumoniae to human type–11 pneumocytes and vascular endothelial cells in vitro" Micro. Path. vol. 17, pp. 361–374 (1994).

Cundell et al., "Streptococcus pneumoniae anchor to activated human cells by the receptor for platelet–activating factor", Nature, vol. 377, pp. 435–438 (1995).

Tuomanen et al., "Alcohol Consumption and Mortality Among Women", New Engl. J. Med., vol. 322, pp. 1280–1284 (1995).

Idanpaan–Heikkila et al., "Oligosaccharides Interfere with the Establishment and Progression of Experimental Pneumococcal Pneumonia", J. Inf. Dis., vol. 176, pp. 704–712 (1997).

Lupas et al., "Predicting Coiled Coils from Protein Sequences", Sciences, vol. 252, pp. 1162–1164 (1991).

* cited by examiner

Primary Examiner—James Housel
Assistant Examiner—Ulrike Winkler
(74) Attorney, Agent, or Firm—Elliot M. Olstein; Alan J. Grant

(57) **ABSTRACT**

A vaccine composition is disclosed that comprises polypeptides and fragments of polypeptides containing histidine triad residues or coiled-coil regions, some of which polypeptides or fragments lie between 80 and 680 residues in length. Also disclosed are processes for preventing infection caused by S. pneumoniae comprising administering of vaccine compositions.

11 Claims, 32 Drawing Sheets

WEST Search History

| Hide Items | Restore | Clear | Cancel |

updated search

DATE: Monday, June 20, 2005

Hide?	Set Name	Query	Hit Count
		DB=USPT; PLUR=YES; OP=AND	
☐	L1	6582706.pn.	1
		DB=PGPB,USPT,USOC,EPAB,JPAB,DWPI,TDBD; PLUR=YES; OP=AND	
☐	L2	lxxc	16
☐	L3	hxxhxh	6
☐	L4	htp$2 or pht$2 or phpa or php-a or bvh$5 or (histidine near2 (motif or domain or triad)) or bvh3 or bvh-3 or pvh3 or pvh-3 or phtb or pht or phtd or phte or pht	14614
☐	L5	L4 and strepto$	1332
☐	L6	L5 and pneumon$	400
☐	L7	L4 same strepto$	69
☐	L8	L7 and pneumon$	43
☐	L9	L7 same pneumon$	34
		DB=EPAB; PLUR=YES; OP=AND	
☐	L10	WO-200288178-A2.did.	0

END OF SEARCH HISTORY

Generate Collection | **Print**

Search Results - Record(s) 1 through 43 of 43 returned.

☐ 1. 20040242844. 16 Jul 04. 02 Dec 04. Group b streptococcus antigens and corresponding dna fragments. Martin, Denis, et al. 530/350; C07K014/00 C07K001/00 C07K017/00.

☐ 2. 20040241687. 18 Jun 04. 02 Dec 04. Novel compounds. Thonnard, Joelle, et al. 435/6; 424/190.1 435/252.3 435/320.1 435/69.3 530/350 536/23.7 C12Q001/68 C07H021/04 A61K039/02 C07K014/195 C12N001/21.

☐ 3. 20040203093. 03 Jun 02. 14 Oct 04. NUCLEOTIDE SEQUENCE OF THE HAEMOPHILUS INFLUENZAE RD GENOME, FRAGMENTS THEREOF, AND USES THEREOF. Fleischmann, Robert D., et al. 435/69.1; 435/320.1 435/325 435/6 536/23.2 702/20 C12Q001/68 G06F019/00 G01N033/48 G01N033/50 C07H021/04.

☐ 4. 20040171113. 19 Apr 04. 02 Sep 04. Antigens of group b streptococcus and corresponding dna fragments. Martin, Denis, et al. 435/69.1; 435/252.3 435/320.1 530/350 536/23.7 C07H021/04 C12N001/21 C07K014/315.

☐ 5. 20040110181. 05 Jan 04. 10 Jun 04. Novel streptococcus pneumoniae open reading frames encoding polypeptide antigens and uses thereof. Zagursky, Robert John, et al. 435/6; 435/252.3 435/320.1 435/69.1 530/350 536/23.7 C12Q001/68 C07H021/04 C07K014/315.

☐ 6. 20040097706. 18 Nov 03. 20 May 04. Streptococcus pyogenes antigens and corresponding dna fragments. Martin, Denis, et al. 530/350; 424/190.1 C07K014/315 C12Q001/68 A61K039/02.

☐ 7. 20040081662. 08 Oct 03. 29 Apr 04. Vaccine. Hermand, Philippe, et al. 424/190.1; A61K039/02.

☐ 8. 20040071730. 26 Sep 03. 15 Apr 04. Bvh-a2 and bvh-a3 antigens of group b streptococcus. Martin, Denis, et al. 424/190.1; 435/252.3 435/320.1 435/6 435/69.1 435/7.32 530/350 C07K014/315 C12N015/74 G01N033/554 G01N033/569 C12Q001/68 A61K039/02.

☐ 9. 20040052820. 08 Oct 02. 18 Mar 04. Fusion proteins comprising DP-178 and other viral fusion inhibitor peptides useful for treating aids. Bolognesi, Dani Paul, et al. 424/208.1; 424/188.1 424/204.1 530/300 530/350 A61K039/21 C07K014/16 A61K039/12 C07K002/00 C07K004/00 C07K005/00 C07K007/00 C07K014/00 C07K016/00 C07K017/00 A61K038/00 C07K001/00.

☐ 10. 20040052781. 14 Apr 03. 18 Mar 04. Vaccine compositions comprising Streptococcus pneumoniae polypeptides having selected structural motifs. Johnson, Leslie S., et al. 424/130.1; 424/185.1 435/100 A61K039/395 A61K039/00.

☐ 11. 20040033235. 06 Jan 03. 19 Feb 04. Nucleic acids encoding DP-178 and other viral fusion inhibitor peptides useful for treating aids. Bolognesi, Dani Paul, et al. 424/186.1; 424/187.1 424/188.1 · 424/208.1 530/350 A61K039/21 A61K039/12 C07K014/16 C07K014/10 C07K014/05 C07K014/11.

☐ 12. 20040005331. 13 Mar 03. 08 Jan 04. Vaccine compositions comprising Streptococcus pneumoniae polypeptides having selected structural motifs. Johnson, Leslie S., et al. 424/190.1; 530/350 536/23.7 A61K039/02 C07H021/04 C07K014/315.

☐   13. 20040001836. 14 Apr 03. 01 Jan 04. Vaccine compositions comprising streptococcus pneumoniae polypeptides having selected structural motifs. Johnson, Leslie S., et al. 424/165.1; 424/190.1 A61K039/40 A61K039/02.

☐   14. 20030204074. 04 Jun 03. 30 Oct 03. Streptococcus pneumoniae 37-kDa surface adhesin a protein. Sampson, Jacquelyn, et al. 536/23.7; 424/190.1 435/252.3 435/320.1 435/69.3 435/7.32 530/350 G01N033/554 G01N033/569 C07H021/04 A61K039/02 C12P021/02 C12N001/21 C07K014/315 C12N015/74.

☐   15. 20030165528. 22 Apr 03. 04 Sep 03. Streptococcus pyogenes antigens. Martin, Denis, et al. 424/190.1; 435/252.3 435/320.1 435/69.3 530/350 536/23.7 A61K039/02 C07H021/04 C07K014/195 C12P021/02 C12N001/21 C12N015/74.

☐   16. 20030105307. 03 Jan 01. 05 Jun 03. Streptococcus pneumoniale 37-kDa surface adhesion a protein. Sampson, Jacquelyn, et al. 536/23.1; 530/350 C07H021/02 C07H021/04 C07K001/00 C07K014/00 C07K017/00.

☐   17. 20030049271. 21 Feb 02. 13 Mar 03. Streptococcus pyogenes polypeptides and corresponding DNA fragments. Martin, Denis, et al. 424/190.1; 435/252.3 435/320.1 435/69.3 435/7.32 530/350 536/23.7 A61K039/02 G01N033/554 G01N033/569 C07K014/315 C07H021/04 C12P021/02 C12N001/21.

☐   18. 20020197605. 18 Dec 00. 26 Dec 02. Novel Polynucleotides. Nakagawa, Satoshi, et al. 435/6; 435/287.2 435/91.2 C12Q001/68 C12P019/34 C12M001/34.

☐   19. 20020032323. 30 Oct 97. 14 Mar 02. STREPTOCOCCUS PNEUMONIAE POLYNUCLEOTIDES AND SEQUENCES. KUNSCH, CHARLES A., et al. 536/23.7; 435/252.3 435/320.1 435/69.1 536/24.32 C12P021/02 C07H021/04 C12N001/21 C12N015/74.

☐   20. 6903184. 10 Jul 00; 07 Jun 05. Multiple antigenic peptides immunogenic against Streptococcus pneumonia. Ades; Edwin W., et al. 530/300; 424/184.1 424/190.1 424/234.1 424/244.1 530/323 530/331 530/350 530/806 530/825. C07K00200 C07K00500 C07K00100 A61K03909 A61K03902.

☐   21. 6833356. 25 Aug 00; 21 Dec 04. Pneumococcal protein homologs and fragments for vaccines. Koenig; Scott, et al. 514/12; 424/130.1 424/184.1 424/243.1 424/244.1 514/2 530/350 536/23.1. C07K014/00 A61K038/16.

☐   22. 6824783. 07 Jun 95; 30 Nov 04. Methods for inhibition of membrane fusion-associated events, including HIV transmission. Bolognesi; Dani Paul, et al. 424/188.1; 424/208.1 435/5 530/360. A61K039/21.

☐   23. 6800744. 30 Jun 98; 05 Oct 04. Nucleic acid and amino acid sequences relating to Streptococcus pneumoniae for diagnostics and therapeutics. Doucette-Stamm; Lynn A., et al. 536/23.1; 435/320.1 435/325 435/419 435/6 536/23.4 536/24.1 536/24.32. C12Q001/68 C12N001/14 C12N015/00 C12N005/00 C12N005/04 C07H021/02 C07H021/04.

☐   24. 6783930. 02 Dec 99; 31 Aug 04. Development of novel anti-microbial agents based on bacteriophage genomics. Pelletier; Jerry, et al. 435/5; 435/7.1 435/7.33 435/7.8 435/883 530/350 530/820 536/23.7. C12Q001/68 G01N033/569 C07H021/04 C07K014/00.

25. <u>6773880</u>. 03 Jan 01; 10 Aug 04. Streptococcus <u>pneumoniale</u> 37-kDa surface adhesion A protein. Sampson; Jacquelyn, et al. 435/6; 536/23.7 536/24.32 536/24.33. C12Q001/68.

26. <u>6617156</u>. 13 Aug 98; 09 Sep 03. Nucleic acid and amino acid sequences relating to Enterococcus faecalis for diagnostics and therapeutics. Doucette-Stamm; Lynn A., et al. 435/320.1; 435/252.3 435/6 435/69.1 536/23.7 536/24.32. C12N015/31 C12N015/63 C12N001/13 C12Q001/68.

27. <u>6583275</u>. 30 Jun 98; 24 Jun 03. Nucleic acid sequences and expression system relating to Enterococcus faecium for diagnostics and therapeutics. Doucette-Stamm; Lynn A., et al. 536/23.1; 435/243 435/320.1 435/325 435/6 536/24.3 536/24.32. C07H021/00 C12Q001/68 C12N015/00 C12N001/00 C12N005/00.

28. <u>6582706</u>. 21 Dec 99; 24 Jun 03. Vaccine compositions comprising Streptococcus <u>pneumoniae</u> polypeptides having selected structural MOTIFS. Johnson; Leslie S., et al. 424/244.1; 424/184.1 424/185.1 424/190.1 424/237.1 435/320.1 435/69.1 530/350 536/23.1 536/23.7. A61K039/09.

29. <u>6562958</u>. 04 Jun 99; 13 May 03. Nucleic acid and amino acid sequences relating to Acinetobacter baumannii for diagnostics and therapeutics. Breton; Gary, et al. 536/23.7; 536/23.1. C07H021/02.

30. <u>6551795</u>. 18 Feb 99; 22 Apr 03. Nucleic acid and amino acid sequences relating to pseudomonas aeruginosa for diagnostics and therapeutics. Rubenfield; Marc J., et al. 435/69.1; 435/253.3 435/320.1 435/325 435/6 536/23.1 536/23.7. C12P021/06 C12N015/00 C07H021/04.

31. <u>6518013</u>. 07 Jun 95; 11 Feb 03. Methods for the inhibition of epstein-barr virus transmission employing anti-viral peptides capable of abrogating viral fusion and transmission. Barney; Shawn O'Lin, et al. 435/5; 424/230.1 530/300 530/324 530/325 530/326. C12Q001/70.

32. <u>6479055</u>. 06 Jun 95; 12 Nov 02. Methods for inhibition of membrane fusion-associated events, including respiratory syncytial virus transmission. Bolognesi; Dani Paul, et al. 424/211.1; 424/186.1 530/324. A61K039/145.

33. <u>6228983</u>. 07 Jun 95; 08 May 01. Human respiratory syncytial virus peptides with antifusogenic and antiviral activities. Barney; Shawn O'Lin, et al. 530/300; 424/186.1 424/211.1 530/324 530/325 530/326. A61K038/00.

34. <u>6217884</u>. 28 Dec 98; 17 Apr 01. Streptococcus <u>pneumoniae</u> 37-kDa surface adhesin a protein. Sampson; Jacquelyn S., et al. 424/244.1; 424/184.1 424/190.1 424/200.1 435/69.1 435/69.3 435/71.1 530/350 536/23.7. A61K039/09.

35. <u>6174860</u>. 16 Apr 99; 16 Jan 01. Insecticidal toxins and nucleic acid sequences coding therefor. Kramer; Vance Cary, et al. 514/12; 435/320.1 435/410 435/412 435/418 435/69.1 536/23.7 800/302. A61K038/00 C12N005/02 C12N005/04 C12P021/06.

36. <u>5874088</u>. 05 Jan 95; 23 Feb 99. Deletion mutants of cholera vaccines expressing heterologous antigens. Mekalanos; John J.. 424/200.1; 424/203.1 424/235.1 424/261.1 435/243 435/252.1 435/252.3 435/69.3 435/909. A61K039/106 C12N001/21.

37. <u>5854416</u>. 17 Sep 96; 29 Dec 98. Streptococcus <u>pneumoniae</u> 37-KDA surface adhesin a protein and nucleic acids coding therefor. Sampson; Jacquelyn S., et al. 536/23.7; 424/244.1 435/320.1

536/23.1. C07H021/04.

☐ 38. <u>WO 200288178A</u>. New <u>BVH-A4</u> proteins and genes from serotype III Group B <u>streptococcus,</u> useful for treating or preventing <u>streptococcal</u> infection in infants, pregnant women, non-pregnant adults (e.g. <u>pneumonia</u>), or members of dairy herd (mastitis). BOYER, M, et al. A61K038/00 A61K039/00 A61K039/09 A61P031/04 C07H021/04 C07K014/315 C07K019/00 C12N001/15 C12N001/19 C12N001/21 C12N005/10 C12N015/09 C12N015/31 C12N015/63 C12P021/02 G01N033/53 G01N033/569 G01N033/68.

☐ 39. <u>WO 200231156A</u>. New <u>BVH-A2 and BVH-A3</u> antigens of Group B <u>Streptococcus,</u> useful for treating, preventing or diagnosing <u>streptococcal</u> infections, e.g. sepsis, meningitis, <u>pneumonia,</u> cellulitis, osteomyelitis, septic arthritis and endocarditis. BOYER, M, et al. A61K038/00 A61K039/02 A61K039/09 A61P031/04 C07K014/315 C07K019/00 C12N001/15 C12N001/19 C12N001/21 C12N005/10 C12N015/09 C12N015/31 C12N015/63 C12N015/74 C12P021/02 C12Q001/68 G01N033/554 G01N033/569.

☐ 40. <u>WO 200222167A</u>. Novel immunogenic composition comprising Streptococcus <u>pneumoniae</u> polysaccharide and protein antigen useful for preventing, ameliorating and treating pneumococcal infections in infants, toddlers and elderly persons. LAFERRIERE, C A J, et al. A61K039/09 A61K039/385 A61K039/39 A61P011/00 A61P027/16 A61P031/04.

☐ 41. <u>WO 200222168A</u>. New immunogenic composition for treating <u>streptococcal</u> infections in infants and elders, comprises two <u>Streptococcus pneumoniae</u> proteins selected from the poly <u>histidine triad</u> family and the choline binding protein family. HERMAND, P, et al. A61K000/00 A61K039/02 A61K039/09 A61K039/385 A61P011/00 A61P027/16 A61P031/04.

☐ 42. <u>WO 200198334A</u>. New <u>Streptococcus pneumoniae BVH-3 and BVH-11</u> variant and epitope-bearing polypeptides, useful as vaccine components for treating or preventing <u>streptococcal</u> infections such as otitis media, meningitis, and bacteremia. BRODEUR, B R, et al. A61K039/00 A61K039/02 A61K039/09 A61K039/39 A61P011/00 A61P027/16 A61P031/04 C07H021/04 C07K014/315 C07K019/00 C12N001/15 C12N001/19 C12N001/21 C12N005/10 C12N009/00 C12N015/09 C12N015/30 C12N015/31 C12N015/63 C12N015/74 C12P021/02 C12N015/63 C12R001:46.

☐ 43. <u>WO 200037105A</u>. Vaccine useful for prophylaxis and treatment of pnemococcal infections such as otitis media, nasopharyngeal and bronchial infections, comprises Streptococcus <u>pneumoniae</u> proteins. ADAMOU, J E, et al. A61K038/00 A61K039/00 A61K039/02 A61K039/09 A61K039/395 A61K039/40 A61P031/04 A61P031/10 C07H021/04 C07K014/315 C07K014:315.

[Generate Collection] [Print]

Terms	Documents
L7 and pneumon$	43

<u>Prev Page</u>   <u>Next Page</u>   <u>Go to Doc#</u>

(12) **United States Patent**
Koenig et al.

(10) **Patent No.:** **US 6,833,356 B1**
(45) **Date of Patent:** **Dec. 21, 2004**

(54) **PNEUMOCOCCAL PROTEIN HOMOLOGS AND FRAGMENTS FOR VACCINES**

(75) Inventors: **Scott Koenig**, Rockville, MD (US); **Jon Heinrichs**, North Potomac, MD (US); **Leslie S. Johnson**, Germantown, MD (US); **John E. Adamou**, Germantown, MD (US)

(73) Assignee: **Medimmune, Inc.**, Gaithersburg, MD (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 197 days.

(21) Appl. No.: **09/645,835**

(22) Filed: **Aug. 25, 2000**

Related U.S. Application Data

(60) Provisional application No. 60/150,750, filed on Aug. 25, 1999.

(51) Int. Cl.[7] **C07K 14/00; A61K 38/16**
(52) U.S. Cl. **514/12; 514/2; 530/350; 424/184.1; 424/130.1; 424/243.1; 424/244.1; 536/23.1**
(58) Field of Search 514/12, 2; 530/350, 530/23.1; 424/184.1, 130.1, 243.1, 244.1, 185.1; 536/23.1

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,694,073 A * 9/1987 Bentle et al. 530/399

2003/0031682 A1 * 2/2003 Brodeur et al. 424/190.1

FOREIGN PATENT DOCUMENTS

WO	WO 98/18930	5/1998
WO	WO 99/42588	8/1999
WO	WO 00/06736	2/2000

OTHER PUBLICATIONS

Spellerberg et al., Lmb, a protein with similarities to the Lral adhesin family, mediates attachment of streptococcus agalactiae to human laminin. Infection and Immunity Feb. 1999, vol. 67 871–878.*

* cited by examiner

Primary Examiner—Robert A. Wax
Assistant Examiner—Chih-Min Kam
(74) *Attorney, Agent, or Firm*—Elliott M. Olstein; Alan J. Grant

(57) **ABSTRACT**

The invention is directed to isolated polypeptides bearing sequence homology to the Sp36 protein found in pneumococcal organisms, such as *Streptococcus pneumoniae*. Polynucleotides encoding such polypeptides are also disclosed. The invention also relates to antibodies specific for the disclosed polypeptides and to uses of such antibodies in the treatment of diseases caused by staphylococci as well as group A and B streptococci. In addition, the invention relates to the use of the disclosed polypeptides in compositions and as vaccines and for prophylactic uses such as in vaccination of animals, especially humans, against a wide variety of streptococcal, staphylococcal and other diseases.

8 Claims, 9 Drawing Sheets
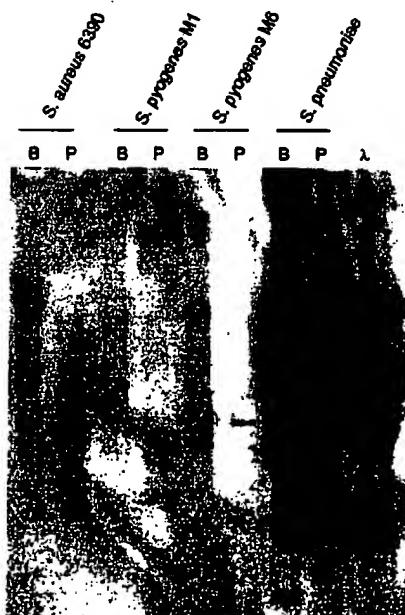
US-PAT-NO: 6833356
DOCUMENT-IDENTIFIER: US 6833356 B1

TITLE: Pneumococcal protein homologs and fragments for vaccines

DATE-ISSUED: December 21, 2004

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Koenig; Scott	Rockville	MD		
Heinrichs; Jon	North Potomac	MD		
Johnson; Leslie S.	Germantown	MD		
Adamou; John E.	Germantown	MD		

ASSIGNEE-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY	TYPE CODE
Medimmune, Inc.	Gaithersburg	MD			02

APPL-NO: 09/ 645835   [PALM]
DATE FILED: August 25, 2000

PARENT-CASE:
This application claims the priority of U.S. Provisional Application No. 60/150,750, filed Aug. 25, 1999, the disclosure of which is hereby incorporated by reference in its entirety.

INT-CL: [07] C07 K 14/00, A61 K 38/16

US-CL-ISSUED: 514/12; 514/2, 530/350, 424/184.1, 424/130.1, 424/243.1, 424/244.1, 536/23.1
US-CL-CURRENT: 514/12; 424/130.1, 424/184.1, 424/243.1, 424/244.1, 514/2, 530/350, 536/23.1

FIELD-OF-SEARCH: 514/12, 514/2, 530/350, 530/23.1, 424/184.1, 424/130.1, 424/243.1, 424/244.1, 424/185.1, 536/23.1

PRIOR-ART-DISCLOSED:

U.S. PATENT DOCUMENTS

| Search Selected | Search ALL | Clear |

PAT-NO	ISSUE-DATE	PATENTEE-NAME	US-CL
4694073	September 1987	Bentle et al.	530/399
2003/0031682	February 2003	Brodeur et al.	424/190.1

FOREIGN PATENT DOCUMENTS

FOREIGN-PAT-NO	PUBN-DATE	COUNTRY	US-CL
WO 98/18930	May 1998	WO	
WO 99/42588	August 1999	WO	
WO 00/06736	February 2000	WO	

OTHER PUBLICATIONS

Spellerberg et al., Lmb, a protein with similarities to the Lral adhesin family, mediates attachment of streptococcus agalactiae to human laminin. Infection and Immunity Feb. 1999, vol. 67 871-878.

ART-UNIT: 1653

PRIMARY-EXAMINER: Wax; Robert A.

ASSISTANT-EXAMINER: Kam; Chih-Min

ATTY-AGENT-FIRM: Olstein; Elliott M. Grant; Alan J.

ABSTRACT:

The invention is directed to isolated polypeptides bearing sequence homology to the Sp36 protein found in pneumococcal organisms, such as Streptococcus pneumoniae. Polynucleotides encoding such polypeptides are also disclosed. The invention also relates to antibodies specific for the disclosed polypeptides and to uses of such antibodies in the treatment of diseases caused by staphylococci as well as group A and B streptococci. In addition, the invention relates to the use of the disclosed polypeptides in compositions and as vaccines and for prophylactic uses such as in vaccination of animals, especially humans, against a wide variety of streptococcal, staphylococcal and other diseases.

8 Claims, 7 Drawing figures

: Entry 15 of 34   File: USPT   Dec 21, 2004

US-PAT-NO: 6833356
DOCUMENT-IDENTIFIER: US 6833356 B1

TITLE: Pneumococcal protein homologs and fragments for vaccines

DATE-ISSUED: December 21, 2004

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Koenig; Scott	Rockville	MD		
Heinrichs; Jon	North Potomac	MD		
Johnson; Leslie S.	Germantown	MD		
Adamou; John E.	Germantown	MD		

US-CL-CURRENT: 514/12; 424/130.1, 424/184.1, 424/243.1, 424/244.1, 514/2, 530/350, 536/23.1

CLAIMS:

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence with at least 95% sequence identity to the sequence of SEQ ID NO: 4 and wherein said polypeptide binds to an antibody that is specific for Sp36 (SEQ ID NO: 7).

2. An isolated polypeptide comprising an amino acid sequence with at least 95% sequence identity to a sequence selected from the group consisting of SEQ ID NO: 2 and 4 wherein said polypeptide is identical to that found in an organism selected from the group consisting of Group A streptococci and Staphylococcus aureus and wherein said polypeptide binds to an antibody that is specific for Sp36 (SEQ ID NO: 7).

3. The isolated polypeptide of claim 2 wherein said Group A organism is Streptococcus pyogenes.

4. The isolated polypeptide of claim 2 wherein said organism is Staphylococcus aureus.

5. An isolated polypepbde comprising an amino acid sequence at least 95% identical to the sequence of SEQ ID NO: 4 and wherein said polypeptide has a sequence with at least 12.6% sequence identity to the amino acid sequence of the Sp36 protein (SEQ ID NO: 7) of Streptococcus pneumoniae and wherein said isolated polypeptide binds to an antibody that is specific for Sp36.

6. An isolated polypeptide comprising the sequence of SEQ ID NO: 2 wherein said isolated polypeptide binds to an antibody that is specific for Sp36 (SEQ ID NO: 7) of Streptococcus pneumoniae.

7. An isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 2.

8. An isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 4.

☐ ▓▓▓ Generate Collection ▓▓▓ ▓ Print ▓

```
    L9: Entry 29 of 34                    File: DWPI            Feb 10, 2005
```

DERWENT-ACC-NO: 2003-120461
DERWENT-WEEK: 200511
COPYRIGHT 2005 DERWENT INFORMATION LTD

TITLE: New BVH-A4 proteins and genes from serotype III Group B streptococcus, useful for treating or preventing streptococcal infection in infants, pregnant women, non-pregnant adults (e.g. pneumonia), or members of dairy herd (mastitis)

INVENTOR: BOYER, M; BRODEUR, B R ; HAMEL, J ; MARTIN, D ; RIOUX, S

PATENT-ASSIGNEE: SHIRE BIOCHEM INC (SHIRN), BOYER M (BOYEI), BRODEUR B R (BRODI), HAMEL J (HAMEI), MARTIN D (MARTI), RIOUX S (RIOUI)

PRIORITY-DATA: 2001US-287712P (May 2, 2001)

▓▓ Search Selected ▓▓   ▓▓ Search ALL ▓▓   ▓▓ Clear ▓▓

PATENT-FAMILY:

	PUB-NO	PUB-DATE	LANGUAGE	PAGES	MAIN-IPC
☐	JP 2005503774 W	February 10, 2005		095	C12N015/09
☐	WO 200288178 A2	November 7, 2002	E	059	C07K014/315
☐	EP 1390505 A2	February 25, 2004	E	000	C12N015/31
☐	AU 2002308325 A1	November 11, 2002		000	C07K014/315
☐	US 20040171113 A1	September 2, 2004		000	C07H021/04

DESIGNATED-STATES: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TN TR TT TZ UA UG US UZ VN YU ZA ZM ZW AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TR TZ UG ZM ZW AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT RO SE SI TR

APPLICATION-DATA:

PUB-NO	APPL-DATE	APPL-NO	DESCRIPTOR
JP2005503774W	May 2, 2002	2002JP-0585476	
JP2005503774W	May 2, 2002	2002WO-CA00664	
JP2005503774W		WO 200288178	Based on
WO 200288178A2	May 2, 2002	2002WO-CA00664	
EP 1390505A2	May 2, 2002	2002EP-0766595	
EP 1390505A2	May 2, 2002	2002WO-CA00664	
EP 1390505A2		WO 200288178	Based on
AU2002308325A1	May 2, 2002	2002AU-0308325	

AU2002308325A1   WO 200288178   Based on
US20040171113A1   May 2, 2002   2002WO-CA00664
US20040171113A1   April 19, 2004   2004US-0476614

INT-CL (IPC): A61 K 38/00; A61 K 39/00; A61 K 39/09; A61 P 31/04; C07 H 21/04; C07
K 14/315; C07 K 19/00; C12 N 1/15; C12 N 1/19; C12 N 1/21; C12 N 5/10; C12 N 15/09;
C12 N 15/31; C12 N 15/63; C12 P 21/02; G01 N 33/53; G01 N 33/569; G01 N 33/68

ABSTRACTED-PUB-NO: WO 200288178A
BASIC-ABSTRACT:

NOVELTY - An isolated polypeptide (designated GBS-BVH-A4), which comprises a BVH-A4
protein polypeptide from serotype III Group B streptococcus (GBS) strain COH1, is
new.

DETAILED DESCRIPTION - An isolated polypeptide (designated GBS-BVH-A4), which
comprises a BVH-A4 protein polypeptide from serotype III Group B streptococcus
(GBS) strain COH1 comprises:

(a) a polypeptide comprising 1055 amino acids (I) fully defined in the
specification;

(b) a polypeptide having at least 80-95% identity to a second polypeptide having
(I);

(c) a polypeptide capable of raising antibodies having binding specificity for (I);

(d) an epitope bearing portion of (a);

(e) fragments or analogs of (a), (b), (c) or (d);

(f) any of the polypeptides of (a)-(e), where the N-terminal Met residue is
deleted; or

(g) any of the polypeptide of (a)-(e), where the secretory amino acid sequence is
deleted.

INDEPENDENT CLAIMS are also included for the following:

(1) An isolated polynucleotide comprising a sequence:

(a) encoding the GBS-BVH-A4 polypeptide;

(b) having 3168 bp fully defined in the specification; or

(c) that is complementary of (a) or (b);

(2) Vectors comprising the polynucleotide, where the DNA is operably linked to an
expression control region;

(3) Host cells comprising the vector;

(4) Producing (M1) the GBS-BVH-A4 polypeptide;

(5) Chimeric polypeptides comprising two or more GBS-BVH-A4 polypeptides, or its fragments or analogs, provided that the polypeptides are linked so as to form a chimeric polypeptide;

(6) A pharmaceutical composition comprising the GBS-BVH-A4 polypeptide, and a pharmaceutical carrier, diluent or adjuvant;

(7) Diagnosing (M2) GBS bacterial infection in a host susceptible to GBS infection; and

(8) A kit comprising the GBS-BVH-A4 polypeptide for detecting or diagnosing streptococcal infection.

ACTIVITY - Antibiotic; Immunostimulant.

Test details are described but no results are given.

MECHANISM OF ACTION - Polypeptide Therapy; Vaccine.

Test details are described but no results are given.

USE - The GBS-BVH-A4 polypeptide is useful for the therapeutic or prophylactic treatment of GBS bacterial infection in a host susceptible to GBS infection. In particular, the GBS-BVH-A4 polypeptide is useful for treating or preventing GBS infection in a neonate or infant (e.g. sepsis, meningitis, pneumonia, cellulitis, osteomyelitis, septic arthritis, endocarditis or epiglotitis), in a pregnant woman (e.g. mild urinary tract infection to life-threatening sepsis and meningitis, osteomyelitis, endocarditis, amniotis, endometritis, wound infection (post-cesarean or post-episiotomy), cellulitis or fasciitis), in a non-pregnant adult (e.g. bacteremia, skin or soft tissue infection, pneumonia, urosepsis, endocarditis, peritonitis, meningitis or emphysema), or in a member of dairy herd (e.g. mastitis). The composition or GBS-BVH-A4 polypeptide is also useful for treating or preventing streptococcal infection (all claimed). The GBS-BVH-A4 polypeptide or antibody is also useful for diagnosing GBS or streptococcal infection.

ABSTRACTED-PUB-NO: WO 200288178A
EQUIVALENT-ABSTRACTS:

CHOSEN-DRAWING: Dwg.0/2

DERWENT-CLASS: B04 D16
CPI-CODES: B04-B04C1; B04-C01G; B04-E01; B04-F0100E; B04-F01E; B04-N0300E; B04-N03E; B11-C07A; B11-C08E; B12-K04A4; B14-A01B2; B14-C03; B14-C09; B14-F01; B14-G01; B14-K01; B14-N01; B14-N07; B14-N16; B14-S06; B14-S11B; D05-C12; D05-H09; D05-H12A; D05-H12B; D05-H12E; D05-H14B; D05-H17A5; D05-H17B5; D05-H17C;

DOCUMENT-IDENTIFIER: US 20040052781 A1
TITLE: Vaccine compositions comprising Streptococcus pneumoniae polypeptides having selected structural motifs

Abstract Paragraph:
A vaccine composition is disclosed that comprises polypeptides and fragments of polypeptides containing histidine triad residues or coiled-coil regions, some of which polypeptides or fragments lie between 80 and 680 residues in length. Also disclosed are processes for preventing infection caused by S. pneumoniae comprising administering of vaccine compositions.

Summary of Invention Paragraph:
[0002] This invention relates generally to the field of bacterial antigens and their use, for example, as immunogenic agents in humans and animals to stimulate an immune response. More specifically, it relates to the vaccination of mammalian species with a polypeptide comprising at least one conserved histidine triad residue (HxxHxH) and at least one helix-forming polypeptide obtained from Streptococcus pneumoniae as a mechanism for stimulating production of antibodies that protect the vaccine recipient against infection by a wide range of serotypes of pathogenic S. pneumoniae. Further, the invention relates to antibodies against such polypeptides useful in diagnosis and passive immune therapy with respect to diagnosing and treating such pneumococcal infections.

Summary of Invention Paragraph:
[0004] Streptococcus pneumoniae is a gram positive bacteria which is a major causative agent in invasive infections in animals and humans, such as sepsis, meningitis, otitis media and lobar pneumonia (Tuomanen et al. New Engl. J. Med. 322:1280-1284 (1995)). As part of the infective process, pneumococci readily bind to non-inflamed human epithelial cells of the upper and lower respiratory tract by binding to eukaryotic carbohydrates in a lectin-like manner (Cundell et al., Micro. Path. 17:361-374 (1994)). Conversion to invasive pneumococcal infections for bound bacteria may involve the local generation of inflammatory factors which may activate the epithelial cells to change the number and type of receptors on their surface (Cundell et al., Nature, 377:435-438 (1995)). Apparently, one such receptor, platelet activating factor (PAF) is engaged by the pneumococcal bacteria and within a very short period of time (minutes) from the appearance of PAF, pneumococci exhibit strongly enhanced adherence and invasion of tissue. Certain soluble receptor analogs have been shown to prevent the progression of pneumococcal infections (Idanpaan-Heikkila et al., J. Inf. Dis., 176:704-712 (1997)). A number of various other proteins have been suggested as being involved in the pathogenicity of S. pneumoniae. There remains a need for identifying polypeptides having epitopes in common from various strains of S. pneumoniae in order to utilize such polypeptides as vaccines to provide protection against a wide variety of S. pneumoniae.

Summary of Invention Paragraph:
[0005] In accordance with the present invention, there is provided vaccines and vaccine compositions that include polypeptides obtained from S. pneumoniae and/or variants of said polypeptides and/or active fragments of such polypeptides.

Brief Description of Drawings Paragraph:
[0013] FIG. 3 is a western blot demonstrating the ability of antisera raised against recombinant Sp36 derived from strain Norway type 4 to react with Sp36 of heterologous strains. Total cell lysates were immunoblotted with mouse antisera to Sp36. A band representing Sp36 protein was detected in all 23 S. pneumoniae strains tested, which included isolates from each of the 23 pneumococcal serotypes represented in the current polysaccharide vaccine.

Brief Description of Drawings Paragraph:

[0018] FIG. 8 shows the results of immunization of mice with PhtD recombinant protein, which leads to protection from lethal sepsis. C3H/HeJ (Panel A and B) or Balb/cByJ (Panel C) mice were immunized subcutaneously with PhtD protein (15 .mu.g in 50 .mu.l PBS emulsified in 50 .mu.l complete Freund's adjuvant (CFA)). The recombinant PhtD protein used in protection experiments consisted of 819 amino acid residues, starting with the cysteine (residue 20). A group of 10 sham-immunized mice received PBS with adjuvant. A second immunization of 15 .mu.g protein with incomplete Freund's adjuvant (IFA) was administered 3 weeks later; the sham group received PBS with IFA. Blood was drawn (retro-orbital bleed) at week 7; and sera from each group was pooled for analysis of anti-PhtD antibody by ELISA. Mice were challenged at week 8 by an intraperitonial (i.p.) injection of approximately 550 CFU S. pneumoniae strain SJ2, serotype 6B (Panel A), 850 CFU of strain EF6796, serotype 6A (Panel B) or 450 CFU of strain EF5668, serotype 4 (Panel C). In preliminary experiments, the LD.sub.50 for strain SJ2 and EF6796 were determined to be approximately 10 CFU for both strains. The LD.sub.50 for strain EF5668 was determined to be <5 CFU. Survival was determined in all groups over the course of 15 days following challenge. Data are presented as the percent survival for a total of 10 mice per experimental group. Two-sample Log-rank test was used for statistical analysis comparing recombinant Pht immunized mice to sham-immunized mice.

Detail Description Paragraph:

[0033] In accordance with a further aspect of the invention, a vaccine of the type hereinabove described is administered for the purpose of preventing or treating infection caused by S. pneumoniae.

Detail Description Paragraph:

[0040] In still another aspect the present invention relates to a method of using one or more antibodies (monoclonal, polyclonal or sera) to the polypeptides of the invention as described above for the prophylaxis and/or treatment of diseases that are caused by pneumococcal bacteria. In particular, the invention relates to a method for the prophylaxis and/or treatment of infectious diseases that are caused by S. pneumoniae. In a still further preferred aspect, the invention relates to a method for the prophylaxis and/or treatment of otitis media, nasopharyngeal, bronchial infections, and the like in humans by utilizing a vaccine of the present invention.

Detail Description Paragraph:

[0054] The present invention further relates to variants of polynucleotides. The variants of the polynucleotides may be a naturally occurring allelic variant of the polynucleotides or a non-naturally occurring variant of the polynucleotides. The variants include variants in which one or more bases are substituted, deleted or inserted. Complements to such coding polynucleotides may be utilized to isolate polynucleotides encoding the same or similar polypeptides. In particular, such procedures are useful to obtain native immunogenic portions of polypeptides from different serotypes of S. pneumoniae, which is especially useful in the production of "chain" polypeptide vaccines containing multiple immunogenic segments.

Detail Description Paragraph:

[0088] The genomic DNA used as target for amplification was isolated from S. pneumoniae Norway strain (serotype 4), the same strain used for genomic sequencing. The complete sequence of the Sp36 gene (SEQ ID NO:9), and its predicted amino acid sequence (SEQ ID NO:8), are given in the Sequence Listing appended hereto. It was noted that the predicted amino acid sequence included a hydrophobic leader sequence followed by a sequence (LSVC) similar to the consensus sequence for Type II signal peptidase (LxxC, in which both x's typically represent small amino acids). Primers (listed as SEQ ID NOS:1-3) were designed that would amplify the Sp36 gene and allow its cloning into pQE10 and expression as a histidine-tagged protein lacking the signal sequence for purification by nickel-affinity chromatography. Cloning of the fragment amplified by SEQ ID Nos 1 and 3 would result in a protein

containing amino acids 2 through 800 of Sp36; cloning of the fragment amplified by SEQ ID Nos 2 and 3 would result in a protein containing amino acids 7 through 800 of Sp36 (amino acid numbers refer to SEQ ID NO:8).

Detail Description Paragraph:
[0090] In each of the three experiments shown in FIGS. 1A-1C, C3H/HeJ mice (10/group) were immunized intraperitoneally (i.p.) with Sp36 protein (15 .mu.g in 50 .mu.l PBS emulsified in 50 .mu.l complete Freund's adjuvant (CFA)). A group of 10 sham-immunized mice received PBS with adjuvant. A second immunization of 15 .mu.g protein with incomplete Freund's adjuvant (IFA) was administered 4 weeks later; the sham group received PBS with IFA. Blood was drawn (retro-orbital bleed) at weeks 3, 6, and 9; and sera from each group were pooled for analysis of anti-Sp36 antibody by ELISA. Mice were challenged at week 10 by an i.p. injection of approximately 500 CFU S. pneumoniae strain SJ2 (serotype 6B; provided by P. Flynn, St. Jude Children's Research Hospital, Memphis, Tenn.). In preliminary experiments, the LD.sub.50 of this strain was determined to be approximately 10 CFU. Mice were monitored for 14 days for survival.

Detail Description Paragraph:
[0099] C3H/HeJ mice (10 mice/group) were passively immunized by two i.p. injections of 100 .mu.l of rabbit serum. The first injection was administered twenty-four hours before challenge with 172 cfu of S. pneumoniae strain SJ2, and the second injection was given four hours after challenge. FIG. 2 shows the survival of mice after infection with two different strains of pneumococci.

Detail Description Paragraph:
[0102] Conservation of Sp36 Among Strains of S. pneumoniae

Detail Description Paragraph:
[0105] The mouse anti-Sp36 sera detected two major bands with apparent molecular weights of 97 and 100 kDa in all 23 pneumococcal lysates tested (shown in FIG. 3). The Sp36 signals obtained from S. pneumoniae serotypes 1, 5, 17F and 22F were lower, indicating either that the level of Sp36 expression is reduced in these strains, or that Sp36 in these strains is antigenically different.

Detail Description Paragraph:
[0112] These experiments indicate that Sp36 is recognized by the human immune system and suggest that antibodies able to bind the Sp36 protein may be produced during natural S. pneumoniae infection in humans. Since the patients were infected with a variety of pneumococcal strains, these data also support the idea that Sp36 is antigenically conserved.

CLAIMS:

2. A process for preventing infection caused by S. pneumoniae comprising: administering the vaccine of claim 1.

☐ ░░░Generate Collection░░░ ░Print░

L9: Entry 10 of 34   File: PGPB   Sep 4, 2003

DOCUMENT-IDENTIFIER: US 20030165528 A1
TITLE: Streptococcus pyogenes antigens

Detail Description Paragraph:
[0110] An ORF which shares 62% with the S. pyogenes BVH-P1 gene was initially
presented in the patent application PCT/CA99/00114 which described Group B
streptococcus antigens. BVH-PL gene was also found to share homology (62% identity)
with an ORF present in the genome of S. pneumoniae (The Institute for Genomic
Research).

DOCUMENT-IDENTIFIER: US 6582706 B1
TITLE: Vaccine compositions comprising Streptococcus pneumoniae polypeptides having selected structural MOTIFS

Brief Summary Text (2):
This invention relates generally to the field of bacterial antigens and their use, for example, as immunogenic agents in humans and animals to stimulate an immune response. More specifically, it relates to the vaccination of mammalian species with a polypeptide comprising at least one conserved histidine triad residue (HxxHxH-SEQ ID NO: 12) and at least one helix-forming polypeptide obtained from Streptococcus pneumoniae as a mechanism for stimulating production of antibodies that protect the vaccine recipient against infection by a wide range of serotypes of pathogenic S. pneumoniae. Further, the invention relates to antibodies against such polypeptides useful in diagnosis and passive immune therapy with respect to diagnosing and treating such pneumococcal infections.

☑	tr	Q8DQ07	_STRR6 Pneumococcal histidine triad protein E [phtE] [S...	2017	
☑	tr	Q6WNQ7	_STRPN Surface protein BVH-3 [bvh-3] [Streptococcus pne...	2017	
☑	tr	Q9ANY1	_STRPN Pneumococcal histidine triad protein E precursor...	2006	
☑	tr	Q6WNQ5	_STRPN Surface protein BVH-3 (Fragment) [bvh-3] [Strept...	1968	
☐	tr	Q8CWR4	_STRR6 Histidine Motif-Containing protein [phpA] [Strep...	442	e-
☑	tr	Q8DPQ2	_STRR6 Pneumococcal histidine triad protein A [phtA] [S...	437	e-
☑	tr	Q9AG74	_STRPN PhpA [phpA] [Streptococcus pneumoniae]	437	e-
☑	tr	Q9AHT9	_STRPN Pneumococcal histidine triad A protein [phtA] [S...	435	e-
☑	tr	Q8DQ08	_STRR6 Pneumococcal histidine triad protein D [phtD] [S...	434	e-
☐	tr	Q6T8D7	_STRPN Pneumococcal protein BVH-11-3 [Streptococcus pne...	431	e-
☐	tr	Q97QM8	_STRPN Conserved domain protein [SP1175] [Streptococcus...	424	e-
☑	tr	Q9ANY2	_STRPN Pneumococcal histidine triad protein D precursor...	423	e-
☐	tr	Q97QM9	_STRPN Conserved domain protein [SP1174] [Streptococcus...	423	e-
☑	tr	Q9ANY3	_STRPN Pneumococcal histidine triad protein B precursor...	423	e-
☐	tr	Q6WNQ3	_STRPN Surface protein BVH-11 (Fragment) [bvh-11] [Stre...	417	e-
☐	tr	Q6WNP8	_STRPN Surface protein BVH-11-2 (Fragment) [bvh-11-2] [...	417	e-
☐	tr	Q6WNQ1	_STRPN Surface protein BVH-11 (Fragment) [bvh-11] [Stre...	415	e-
☐	tr	Q6WNP5	_STRPN Surface protein BVH-11-2 (Fragment) [bvh-11-2] [...	415	e-
☐	tr	Q6WNP9	_STRPN Surface protein BVH-11-2 (Fragment) [Streptococc...	410	e-
☐	tr	Q6T304	_STRPN Surface protein BVH-11 (Fragment) [bvh11] [Strep...	410	e-
☐	tr	Q6WNQ0	_STRPN Surface protein BVH-11 (Fragment) [bvh-11] [Stre...	405	e-
☐	tr	Q6WNP6	_STRPN Surface protein BVH-11-2 (Fragment) [bvh-11-2] [...	400	e-
☐	tr	Q8NZ82	_STRP8 Hypothetical protein spyM18_2072 [spyM18_2072] [...	306	1e
☐	tr	Q8E4U1	_STRA3 Hypothetical protein gbs1306 [gbs1306] [Streptoc...	305	3e
☐	tr	Q8DZ81	_STRA5 Streptococcal histidine triad family protein [SA...	305	3e
☐	tr	Q5X9R2	_STRP6 Streptococcal histidine triad protein [M6_Spy171...	305	4e
☐	tr	Q9ZHG7	_STRAG Hypothetical protein [Streptococcus agalactiae]	305	4e
☐	tr	Q99XV4	_STRPY Hypothetical protein SPy2006 [SPy2006] [Streptoc...	304	6e
☐	tr	Q8K5Q1	_STRP3 Histidine triad protein [SpyM3_1724] [Streptococ...	304	8e
☐	tr	Q93GT5	_STRPY Histidine triad protein of group A streptococci ...	302	2e
☐	tr	Q8E338	_STRA3 Hypothetical protein gbs1925 [gbs1925] [Streptoc...	301	5e
☐	tr	Q877Y2	_STRP3 Hypothetical protein SPs1722 [SPs1722] [Streptoc...	291	4e
☐	tr	Q9AE21	_STRAG Hypothetical protein (Fragment) [Streptococcus a...	241	9e
☐	tr	Q8DQ06	_STRR6 Pneumococcal histidine triad protein E, truncati...	121	8e
☐	tr	Q8E029	_STRA5 Hypothetical protein SAG0907 [SAG0907] [Streptoc...	111	1e
☐	tr	Q8E5R2	_STRA3 Hypothetical protein gbs0918 [gbs0918] [Streptoc...	109	3e
☐	tr	Q8P0G5	_STRP8 Putative internalin A [spyM18_1373] [Streptococc...	99	4e
☐	tr	Q8K714	_STRP3 Putative internalin A [inlA] [Streptococcus pyog...	97	2e
☐	tr	Q5XBJ5	_STRP6 Internalin protein [M6_Spy1083] [Streptococcus p...	97	3e
☐	tr	Q99Z76	_STRPY Putative internalin A [inlA] [Streptococcus pyog...	92	9e
☐	tr	Q5AWR8	_EMENI Hypothetical protein [AN7262.2] [Aspergillus nid...	47	0.
☐	tr	Q6HCJ0	_BACHK Cell division protein [ftsK] [Bacillus thuringie...	47	0.
☐	sp	P45386	IGA4_HAEIN Immunoglobulin A1 protease precursor (EC 3....	46	0.
☐	tr	Q8ISF7	_CAEEL 2MDa_1 protein [isof] [Caenorhabditis elegans]	46	0.

☐	tr	Q8ISF6	_CAEEL 2MDa_2 protein [isof] [Caenorhabditis elegans]	46	0.
☐	sp	P16053	NFM_CHICK Neurofilament triplet M protein (160 kDa neu...	45	0.
☐	sp	Q97QP7	IGA1A_STRPN Immunoglobulin A1 protease precursor (EC 3...	44	0.
☐	tr	Q869E1	_DICDI DNA ligase I (EC 6.5.1.1) (Polydeoxyribonucleoti...	44	0.
☐	tr	Q8IB63	_PLAF7 Hypothetical protein PF08_0035 [PF08_0035] [Plas...	44	0.
☐	tr	Q9VC00	_DROME CG13648-PA [CG13648] [Drosophila melanogaster (F...	44	0.
☐	tr	Q5TUJ9	_ANOGA ENSANGP00000029120 [ENSANGG00000022532] [Anophel...	44	0.
☐	tr	Q839C3	_ENTFA N-acetylmuramoyl-L-alanine amidase, family 4 [EF...	44	0.
☐	tr	Q7PR93	_ANOGA ENSANGP00000010616 [ENSANGG00000008127] [Anophel...	44	0.
☐	tr	O73793	_SERCA Neurofilament medium subunit [Serinus canaria (C...	43	0.
☐	tr	Q9FN97	_ARATH Transposon protein-like [Arabidopsis thaliana (M...	43	0.
☐	tr	Q54U33	_DICDI Hypothetical protein [DDB0218298] [Dictyostelium...	43	0.
☐	tr	Q963T1	_PLARE Glutamate-rich protein (Fragment) [GLURP] [Plasm...	43	0.
☐	tr	Q07594	_ENTHI K2 protein (Fragment) [Entamoeba histolytica]	43	0.
☐	tr	Q6PK21	_HUMAN OGFR protein [Homo sapiens (Human)]	43	0.
☐	sp_vs	Q9NZT2-2	Splice isoform 2 of Q9NZT2 [OGFR] [Homo sapiens (...	43	0.
☐	sp	Q9NZT2	OGFR_HUMAN Opioid growth factor receptor (OGFr) (Zeta-...	43	0.
☐	tr	Q6HBX5	_BACHK Hypothetical protein [BT9727_4640] [Bacillus thu...	43	0.
☐	tr	Q7RQS8	_PLAYO Retinitis pigmentosa GTPase regulator-like prote...	43	0.
☐	tr	O77320	_PLAF7 Hypothetical protein MAL3P3.3 [MAL3P3.3] [Plasmo...	43	0.
☐	tr	Q5HQ74	_STAEQ Pyruvate dehydrogenase complex E2 component, dih...	42	0.
☐	tr	Q9L4Z1	_STAEP Pyruvate dehydrogenase complex subunit E2 [pdhC]...	42	0.
☐	tr	Q9VGW4	_DROME CG14692-PA [CG14692] [Drosophila melanogaster (F...	42	0.
☐	tr	Q90307	_CARAU Carassius auratus [Carassius auratus (Goldfish)]	42	0.
☐	tr	Q898B0	_CLOTE Hypothetical protein CTC00555 [CTC00555] [Clostr...	42	0.
☐	tr	Q6R4Z8	_CAPBU Dehydrin cor29 [Capsella bursa-pastoris (Shepher...	42	0.
☐	tr	Q50R78	_ENTHI Hypothetical protein [298.t00012] [Entamoeba his...	42	(
☐	tr	Q8MMQ1	_DICDI Similar to Required for the transfer of mannosyl...	42	(
☐	tr	Q8IBL1	_PLAF7 Hypothetical protein MAL7P1.129 [MAL7P1.129] [Pl...	42	(
☐	tr	Q55K21	_CRYNE Hypothetical protein [CNBK1650] [Cryptococcus ne...	42	(
☐	sp	Q8CT13	ODP2_STAEP Dihydrolipoyllysine-residue acetyltransfera...	41	(
☐	tr	O33741	_STRPN SpsA protein [Streptococcus pneumoniae]	41	(
☐	tr	Q8I1W5	_PLAF7 Hypothetical protein PFD0320c (Fragment) [PFD032...	41	(
☐	tr	Q5TVN3	_ANOGA ENSANGP00000027660 (Fragment) [ENSANGG0000002494...	41	(
☐	tr	Q07593	_ENTHI K2 protein (Fragment) [Entamoeba histolytica]	41	(
☐	sp	Q28820	TRDN_RABIT Triadin [TRDN] [Oryctolagus cuniculus (Rabb...	41	(
☐	sp	Q59947	IGA1_STRR6 Immunoglobulin A1 protease precursor (EC 3....	41	(
☐	tr	Q7SXW9	_BRARE Wu:fc44a10 protein (Fragment) [wu:fc44a10] [Brac...	41	(
☐	tr	Q28688	_RABIT Neurofilament-H (Fragment) [Oryctolagus cuniculu...	41	(
☐	tr	Q94674	_PLAGA Thrombospondin-related anonymous protein (Fragme...	41	(
☐	tr	Q5WNG8	_CAEBR Hypothetical protein CBG08011 [CBG08011] [Caenor...	41	(
☐	tr	Q6BLN0	_DEBHA Similar to ca|CA2433|IPF12959 Candida albicans I...	41	(
☐	sp_vs	Q28820-4	Splice isoform Cardiac 3 of Q28820 [TRDN] [Orycto...	41	(
☐	sp_vs	Q28820-6	Splice isoform Skeletal 3 of Q28820 [TRDN] [Oryct...	41	(
☐	sp	Q54875	IGA1B_STRPN Immunoglobulin A1 protease precursor (EC 3...	40	(

☐ tr	Q568L0	_BRARE	Wu:fc44a10 [wu:fc44a10] [Brachydanio rerio (Zebr...	40	(
☐ tr	Q4ZHU3	_STAXY	Biofilm-associated protein [bap] [Staphylococcus...	40	(
☐ tr	Q510B7	_ENTHI	Hypothetical protein [113.t00023] [Entamoeba his...	40	(
☐ tr	Q8IJ56	_PLAF7	Glutamate-rich protein [PF10_0344] [Plasmodium f...	40	(
☐ tr	Q9GUY4	_PENJP	Crustocalcin [DD4(ccn)] [Penaeus japonicus (Kuru...	40	(
☐ tr	Q9GTX2	_PLAFA	Glutamate-rich protein [GLURP] [Plasmodium falci...	40	(
☐ tr	Q61US9	_CAEBR	Hypothetical protein CBG05170 [CBG05170] [Caenor...	40	(
☐ tr	Q6FWC0	_CANGA	Candida glabrata strain CBS138 chromosome D comp...	40	(
☐ tr	Q6CTI0	_KLULA	Similar to sp|Q05050 Saccharomyces cerevisiae YM...	40	(
☐ tr	Q879S6	_XYLFT	Hemagglutinin-like secreted protein [pspA] [Xyle...	40	(
☐ tr	O17102	_CAEEL	Hypothetical protein F42G2.6 [F42G2.6] [Caenorha...	40	(

CLUSTAL W (1.74) multiple sequence alignment

```
tr|Q8DQ07|Q8DQ07_STRR6     ------------MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENK-DNN
tr|Q6WNQ7|Q6WNQ7_STRPN     ------------MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENK-DNN
tr|Q9ANY1|Q9ANY1_STRPN     ------------MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENK-DNN
tr|Q6WNQ5|Q6WNQ5_STRPN     ----------------------------CAYALNQHRSQENK-DNN
tr|Q8DPQ2|Q8DPQ2_STRR6     MQLEISNRKRVSMKINKKYLVG-SAAALILSVCSYELGLYQARTVK-ENN
tr|Q9AG74|Q9AG74_STRPN     ------------MKINKKYLVG-SAAALILSVCSYELGLYQARTVK-ENN
tr|Q9AHT9|Q9AHT9_STRPN     ------------MKINKKYLVG-SAAALILSVCSYELGLYQARTVK-ENN
tr|Q8DQ08|Q8DQ08_STRR6     ------------MKINKKYLAG-SVAVLALSVCSYELGRHQAGQVKKESN
tr|Q9ANY2|Q9ANY2_STRPN     ------------MKINKKYLAG-SVAVLALSVCSYELGRHQAGQVKKESN
tr|Q9ANY3|Q9ANY3_STRPN     ------------MKINKKYLAG-SVAVLALSVCSYELGRYQAGQDKKESN
                                       *:* *. ::: * :.*

tr|Q8DQ07|Q8DQ07_STRR6     RVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYVTSHGDHY
tr|Q6WNQ7|Q6WNQ7_STRPN     RVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYVTSHGDHY
tr|Q9ANY1|Q9ANY1_STRPN     RVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYVTSHGDHY
tr|Q6WNQ5|Q6WNQ5_STRPN     RVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYVTSHGDHY
tr|Q8DPQ2|Q8DPQ2_STRR6     RVSYIDGKQATQKTENLTPDEVSKREGINAEQIVIKITDQGYVTSHGDHY
tr|Q9AG74|Q9AG74_STRPN     RVSYIDGKQATQKTENLTPDEVSKREGINAEQIVIKITDQGYVTSHGDHY
tr|Q9AHT9|Q9AHT9_STRPN     RVSYIDGKQATQKTENLTPDEVSKREGINAEQIVIKITDQGYVTSHGDHY
tr|Q8DQ08|Q8DQ08_STRR6     RVSYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGYVTSHGDHY
tr|Q9ANY2|Q9ANY2_STRPN     RVSYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGYVTSHGDHY
tr|Q9ANY3|Q9ANY3_STRPN     RVAYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGYVTSHGDHY
                           **:*:**.*: **:******:**::***:*******************

tr|Q8DQ07|Q8DQ07_STRR6     HYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVY
tr|Q6WNQ7|Q6WNQ7_STRPN     HYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVY
tr|Q9ANY1|Q9ANY1_STRPN     HYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVY
tr|Q6WNQ5|Q6WNQ5_STRPN     HYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVY
tr|Q8DPQ2|Q8DPQ2_STRR6     HYYNGKVPYDAIFSEELLMKDPNYKLKDEDIVNEVKGGYVIKVDGKYYVY
tr|Q9AG74|Q9AG74_STRPN     HYYNGKVPYDAIISEELLMKDPNYQLKDEDIISEIKGGYVIKVDGKYYVY
tr|Q9AHT9|Q9AHT9_STRPN     HYYNGKVPYDAIISEELLMKDPNYKLKDEDIVNEVKGGYVIKVDGKYYVY
tr|Q8DQ08|Q8DQ08_STRR6     HYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVY
tr|Q9ANY2|Q9ANY2_STRPN     HYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVY
tr|Q9ANY3|Q9ANY3_STRPN     HYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYVY
                           **********:*************:*** **:.*:****:***:******

tr|Q8DQ07|Q8DQ07_STRR6     LKDAAHADNVRTKDEINRQKQEHVKDNE----KVNSNVAVARSQGRYTTN
tr|Q6WNQ7|Q6WNQ7_STRPN     LKDAAHADNVRTKDEINRQKQEHVKDNE----KVNSNVAVARSQGRYTTN
tr|Q9ANY1|Q9ANY1_STRPN     LKDAAHADNVRTKDEINRQKQEHVKDNE----KVNSNVAVARSQGRYTTN
tr|Q6WNQ5|Q6WNQ5_STRPN     LKDAAHADNVRTKDEINRQKQEHVKDNE----KVNSNVAVARSQGRYTTN
tr|Q8DPQ2|Q8DPQ2_STRR6     LKDAAHADNVRTKEEINRQKQEHSQHREGGTPRNDGAVALARSQGRYTTD
tr|Q9AG74|Q9AG74_STRPN     LKDAAHADNVRTKEEINRQKQEHSQHREGGTPRNDGAVALARSQGRYTTD
tr|Q9AHT9|Q9AHT9_STRPN     LKDAAHADNVRTKEEINRQKQEHSQHREGGTPRNDGAVALARSQGRYTTD
tr|Q8DQ08|Q8DQ08_STRR6     LKDAAHADNIRTKEEIKRQKQERSHNHN---SRADNAVAAARAQGRYTTD
tr|Q9ANY2|Q9ANY2_STRPN     LKDAAHADNIRTKEEIKRQKQEHSHNHGG--GSNDQAVVAARAQGRYTTD
tr|Q9ANY3|Q9ANY3_STRPN     LKDAAHADNIRTKEEIKRQKQERSHNHN---SRADNAVAAARAQGRYTTD
                           *********:***:**:*****: :..      :  *. **:******:

tr|Q8DQ07|Q8DQ07_STRR6     DGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGK-
tr|Q6WNQ7|Q6WNQ7_STRPN     DGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGK-
tr|Q9ANY1|Q9ANY1_STRPN     DGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGK-
tr|Q6WNQ5|Q6WNQ5_STRPN     DGYVFNPADIIEDTGNAYIVPHRGHYHYIPKSDLSASELAAAKAHLAGK-
tr|Q8DPQ2|Q8DPQ2_STRR6     DGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAFLSGRG
tr|Q9AG74|Q9AG74_STRPN     DGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAKAFLSGRG
tr|Q9AHT9|Q9AHT9_STRPN     DGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAFLSGRG
tr|Q8DQ08|Q8DQ08_STRR6     DGYIFNASDIIEDTGDAYIVPHGDHYHYIPKSDLSASELAAAQAYWNGK-
tr|Q9ANY2|Q9ANY2_STRPN     DGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAYWNGK-
```

```
tr|Q9ANY3|Q9ANY3_STRPN       DGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAYWNGK-
                             ***:**.:*******:****** .*******.:**********:*.  *:


tr|Q8DQ07|Q8DQ07_STRR6       ----------NMQP-SQLSYSSTASD---NNTQSVAKGSTSKPANKSEN
tr|Q6WNQ7|Q6WNQ7_STRPN       ----------NMQP-SQLSYSSTASD---NNTQSVAKGSTSKPANKSEN
tr|Q9ANY1|Q9ANY1_STRPN       ----------NMQP-SQLSYSSTASD---NNTQSVAKGSTSKPANKSEN
tr|Q6WNQ5|Q6WNQ5_STRPN       ----------NMQP-SQLSYSSTASD---NNTQSVAKGSTSKPANKSEN
tr|Q8DPQ2|Q8DPQ2_STRR6       NLSNSRTYRRQNSDNTSRTNWVPSVSNPGTTNTNTSNNSNTNSQASQSND
tr|Q9AG74|Q9AG74_STRPN       NLSNSRTYRRQNSDNTSRTNWVPSVSNPGTTNTNTSNNSNTNSQASQSND
tr|Q9AHT9|Q9AHT9_STRPN       NLSNSRTYRRQNSDNTSRTNWVPSVSNPGTTNTNTSNNSNTNSQASQSND
tr|Q8DQ08|Q8DQ08_STRR6       ----------QGSRPSSSSSHNANPAQPRLSENHNLTVTPTYHQN-QGEN
tr|Q9ANY2|Q9ANY2_STRPN       ----------QGSRPSSSSSYNANPAQPRLSENHNLTVTPTYHQN-QGEN
tr|Q9ANY3|Q9ANY3_STRPN       ----------QGSRPSSSSSYNANPAQPRLSENHNLTVTPTYHQN-QGEN
                                      .    *   .. ::   .:.:.    *      :.::


tr|Q8DQ07|Q8DQ07_STRR6       LQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIP
tr|Q6WNQ7|Q6WNQ7_STRPN       LQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIP
tr|Q9ANY1|Q9ANY1_STRPN       LQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIP
tr|Q6WNQ5|Q6WNQ5_STRPN       LQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIP
tr|Q8DPQ2|Q8DPQ2_STRR6       IDSLLKQLYKLPLSQRHVESDGLVFDPAQITSRTARGVAVPHGDHYHFIP
tr|Q9AG74|Q9AG74_STRPN       IDSLLKQLYKLPLSQRHVESDGLIFDPAQITSRTANGVAVPHGDHYHFIP
tr|Q9AHT9|Q9AHT9_STRPN       IDSLLKQLYKLPLSQRHVESDGLVFDPAQITSRTARGVAVPHGDHYHFIP
tr|Q8DQ08|Q8DQ08_STRR6       ISSLLRELYAKPLSERHVESDGLIFDPAQITSRTANGVAVPHGDHYHFIP
tr|Q9ANY2|Q9ANY2_STRPN       ISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGVAVPHGNHYHFIP
tr|Q9ANY3|Q9ANY3_STRPN       ISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGVAVPHGNHYHFIP
                             :.***::**   *  ::*: *****:****:* ***..***:***:******


tr|Q8DQ07|Q8DQ07_STRR6       YSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVSSLGSLSSNPS---SL
tr|Q6WNQ7|Q6WNQ7_STRPN       YSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVSSLGSLSSNPS---SL
tr|Q9ANY1|Q9ANY1_STRPN       YSKLSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSNPS---SL
tr|Q6WNQ5|Q6WNQ5_STRPN       YSKLSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSNPS---SL
tr|Q8DPQ2|Q8DPQ2_STRR6       YSQMSELEERIARIIPLRYRSNHWVPDSRP-EQPSPQPTPEPSPG-----
tr|Q9AG74|Q9AG74_STRPN       YSQLSPLEEKLARIIPLRYRSNHWVPDSRP-EQPSPQSTPEPSPSPQPAP
tr|Q9AHT9|Q9AHT9_STRPN       YSQMSELEERIARIIPLRYRSNHWVPDSRP-EQPSPQPTPEPSPG-----
tr|Q8DQ08|Q8DQ08_STRR6       YSQLSPLEEKLARIIPLRYRSNHWVPDSRP-EQPSPQSTPEPSPSPQPAP
tr|Q9ANY2|Q9ANY2_STRPN       YEQMSELEKRIARIIPLRYRSNHWVPDSRP-EQPSPQSTPEPSPSPQPAP
tr|Q9ANY3|Q9ANY3_STRPN       YEQMSELEKRIARIIPLRYRSNHWVPDSRP-EEPSPQPTPEPSPS-----
                             *.::* **::**  :*:     ..    .:::* *  *.  :  ...*.


tr|Q8DQ07|Q8DQ07_STRR6       TTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTL
tr|Q6WNQ7|Q6WNQ7_STRPN       TTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTL
tr|Q9ANY1|Q9ANY1_STRPN       TTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTL
tr|Q6WNQ5|Q6WNQ5_STRPN       TTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTL
tr|Q8DPQ2|Q8DPQ2_STRR6       -PQPAPNLKIDSNSSLVSQLVRKVGEGYVFEEKGISRYVFAKD------L
tr|Q9AG74|Q9AG74_STRPN       NPQPAPSNPIDEK--LVKEAVRKVGDGYVFEENGVPRYIPAKD------L
tr|Q9AHT9|Q9AHT9_STRPN       -PQPAPNLKIDSNSSLVSQLVRKVGEGYVFEEKGISRYVFAKD------L
tr|Q8DQ08|Q8DQ08_STRR6       NPQPAPSNPIDEK--LVKEAVRKVGDGYVFEENGVPRYIPAKD------L
tr|Q9ANY2|Q9ANY2_STRPN       NPQPAPSNPIDEK--LVKEAVRKVGDGYVFEENGVSRYIPAKD------L
tr|Q9ANY3|Q9ANY3_STRPN       -PQPAPSNPIDGK--LVKEAVRKVGDGYVFEENGVSRYIPAKD------L
                             ..    *        .: *.:.. .*:... .  :*:  .:       *


tr|Q8DQ07|Q8DQ07_STRR6       PNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNH
tr|Q6WNQ7|Q6WNQ7_STRPN       PNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNH
tr|Q9ANY1|Q9ANY1_STRPN       PNNSLATPSPSLPINPGTSHEKHEEDGYGFDANRIIAEDESGFVMSHGDH
tr|Q6WNQ5|Q6WNQ5_STRPN       PNNSLATPSPSLPINPGTSHEKHEEDGYGFDANRIIAEDESGFVMSHGDH
tr|Q8DPQ2|Q8DPQ2_STRR6       PSETVKNLESKLSKQESVSHT------------LTAKKEN---VAPRDQ
tr|Q9AG74|Q9AG74_STRPN       SAETAAGIDSKLAKQESLSHK------------LGAKKTD---LPSSDR
tr|Q9AHT9|Q9AHT9_STRPN       PSETVKNLESKLSKQESVSHT------------LTAKKEN---VAPRDQ
tr|Q8DQ08|Q8DQ08_STRR6       SAETAAGIDSKLAKQESLSHK------------LGAKKTD---LPSSDR
tr|Q9ANY2|Q9ANY2_STRPN       SAETAAGIDSKLAKQESLSHK------------LGAKKTD---LPSSDR
```

```
tr|Q9ANY3|Q9ANY3_STRPN        SAETAAGIDSKLAKQESLSHK-------------LGTKKTD---LPSSDR
                              .  ::      ...*. : . **              : ::. .    :.  ::


tr|Q8DQ07|Q8DQ07_STRR6        NHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKD
tr|Q6WNQ7|Q6WNQ7_STRPN        NHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKD
tr|Q9ANY1|Q9ANY1_STRPN        NHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPSNAKEMKD
tr|Q6WNQ5|Q6WNQ5_STRPN        NHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPSNAKEMKD
tr|Q8DPQ2|Q8DPQ2_STRR6        EFYDKAYNLLTEAHKALFEN-KGRNSDFQALDKLLERLNDEST------N
tr|Q9AG74|Q9AG74_STRPN        EFYNKAYDLLARIHQDLLDN-KGRQVDFEALDNLLERLKDVSS------D
tr|Q9AHT9|Q9AHT9_STRPN        EFYDKAYNLLTEAHKALFXN-KGRNSDFQALDKLLERLNDEST------N
tr|Q8DQ08|Q8DQ08_STRR6        EFYNKAYDLLARIHQDLLDN-KGRQVDFEALDNLLERLKDVSS------D
tr|Q9ANY2|Q9ANY2_STRPN        EFYNKAYDLLARIHQDLLDN-KGRQVDFEALDNLLERLKDVPS------D
tr|Q9ANY3|Q9ANY3_STRPN        EFYNKAYDLLARIHQDLLDN-KGRQVDFEALDNLLERLKDVSS------D
                              :.*      :*   .  :       : :   :  ..:.**.* .: :* .        :


tr|Q8DQ07|Q8DQ07_STRR6        LDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPIDEHKP
tr|Q6WNQ7|Q6WNQ7_STRPN        LDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPIDEHKP
tr|Q9ANY1|Q9ANY1_STRPN        LDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPIDEHKP
tr|Q6WNQ5|Q6WNQ5_STRPN        LDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPIDEHKP
tr|Q8DPQ2|Q8DPQ2_STRR6        KEKLVDDLLAFLAPITHPER----LGKPNSQIEYTE------DEVRIAQL
tr|Q9AG74|Q9AG74_STRPN        KVKLVDDILAFLAPIRHPER----LGKPNAQITYTD------DEIQVAKL
tr|Q9AHT9|Q9AHT9_STRPN        KEKLVDDLLAFLAPITHPER----LGKPNSQIEYTE------DEVRIAQL
tr|Q8DQ08|Q8DQ08_STRR6        KVKLVDDILAFLAPIRHPER----LGKPNAQITYTD------DEIQVAKL
tr|Q9ANY2|Q9ANY2_STRPN        KVKLVDDILAFLAPIRHPER----LGKPNAQITYTD------DEIQVAKL
tr|Q9ANY3|Q9ANY3_STRPN        KVKLVEDILAFLAPIRHPER----LGKPNAQITYTD------DEIQVAKL
                                * ::: :* :        :*      :.* :   * *..      * :    :


tr|Q8DQ07|Q8DQ07_STRR6        VGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNF
tr|Q6WNQ7|Q6WNQ7_STRPN        VGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNF
tr|Q9ANY1|Q9ANY1_STRPN        VGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNF
tr|Q6WNQ5|Q6WNQ5_STRPN        VGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNF
tr|Q8DPQ2|Q8DPQ2_STRR6        ADK-YTTSDGYIFDEHD-IISDEGD-AYVTPHMGHS-HWIGKDSLSDKEK
tr|Q9AG74|Q9AG74_STRPN        AGK-YTTEDGYIFDPRD-ITSDEGD-AYVTPHMTHS-HWIKKDSLSEAER
tr|Q9AHT9|Q9AHT9_STRPN        ADK-YTTSDGYIFDEHD-IISDEGD-AYVTPHMGHS-HWIGKDSLSDKEK
tr|Q8DQ08|Q8DQ08_STRR6        AGK-YTTEDGYIFDPRD-ITSDEGD-AYVTPHMTHS-HWIKKDSLSEAER
tr|Q9ANY2|Q9ANY2_STRPN        AGK-YTTEDGYIFDPRD-ITSDEGD-AYVTPHMTHS-HWIKKDSLSEAER
tr|Q9ANY3|Q9ANY3_STRPN        AGK-YTAEDGYIFDPRD-ITSDEGD-AYVTPHMTHS-HWIKKDSLSEAER
                              ..   ::  .:   :*. .:  :  ..**: .*.    .: :   : :  :.:*.: :


tr|Q8DQ07|Q8DQ07_STRR6        TLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVG
tr|Q6WNQ7|Q6WNQ7_STRPN        TLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVG
tr|Q9ANY1|Q9ANY1_STRPN        TLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVG
tr|Q6WNQ5|Q6WNQ5_STRPN        TLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVG
tr|Q8DPQ2|Q8DPQ2_STRR6        VAAQAYTKEKGILPPSPDADVKAN-------PTGDSAAAIYNRVKGE---
tr|Q9AG74|Q9AG74_STRPN        AAAQAYAKEKGLTPPSTDHQDSGN-------TEAKGAEAIYNRVKAA---
tr|Q9AHT9|Q9AHT9_STRPN        VAAQAYTKEKGILPPSPDADVKAN-------PTGDSAAAIYNRVKGE---
tr|Q8DQ08|Q8DQ08_STRR6        AAAQAYAKEKGLTPPSTDHQDSGN-------TEAKGAEAIYNRVKAA---
tr|Q9ANY2|Q9ANY2_STRPN        AAAQAYAKEKGLTPPSTDHQDSGN-------TEAKGAEAIYNRVKAA---
tr|Q9ANY3|Q9ANY3_STRPN        AAAQAYAXEKGLTPPSTDHQDSGN-------TEAKGAEAIYNRVKAA---
                              . *:.    .   **.:.   *        . ..    : .:* .


tr|Q8DQ07|Q8DQ07_STRR6        NIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTI
tr|Q6WNQ7|Q6WNQ7_STRPN        NIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTI
tr|Q9ANY1|Q9ANY1_STRPN        NIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTI
tr|Q6WNQ5|Q6WNQ5_STRPN        NIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTI
tr|Q8DPQ2|Q8DPQ2_STRR6        --KRIPLVR--LP-YMVEHTVEVK-------NGNLIIP---HKDHYHNIK
tr|Q9AG74|Q9AG74_STRPN        --KKVPLDR--MP-YNLQYTVEVK-------NGSLIIP---HYDHYHNIK
tr|Q9AHT9|Q9AHT9_STRPN        --KRIPLVR--LP-YMVEHTVEVK-------NGNLIIP---HKDHYHNIK
tr|Q8DQ08|Q8DQ08_STRR6        --KKVPLDR--MP-YNLQYTVEVK-------NGSLIIP---HYDHYHNIK
tr|Q9ANY2|Q9ANY2_STRPN        --KKVPLDR--MP-YNLQYTVEVK-------NGSLIIP---HYDHYHNIK
```

```
tr|Q9ANY3|Q9ANY3_STRPN    --KKVPLDR--MP-YNLQYTVEVK-------NGSLIIP---HYDHYHNIK
                            .. * :  :*    .::*:  *        :*.: :*     .   :


tr|Q8DQ07|Q8DQ07_STRR6    FYPFHAGDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKL
tr|Q6WNQ7|Q6WNQ7_STRPN    FYPFHAGDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKL
tr|Q9ANY1|Q9ANY1_STRPN    FYPFHAGDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKL
tr|Q6WNQ5|Q6WNQ5_STRPN    FYPFHAGDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKL
tr|Q8DPQ2|Q8DPQ2_STRR6    FAWFDDH--------TYKAPNG-YTLEDLFATIK--YYVEHPDER-----
tr|Q9AG74|Q9AG74_STRPN    FEWFDEG--------LYEAPKG-YSLEDLLATVK--YYVEHPNER-----
tr|Q9AHT9|Q9AHT9_STRPN    FAWFDDH--------TYKAPNG-YTLEDLFATIK--YYVEHPDER-----
tr|Q8DQ08|Q8DQ08_STRR6    FEWFDEG--------LYEAPKG-YSLEDLLATVK--YYVEHPNER-----
tr|Q9ANY2|Q9ANY2_STRPN    FEWFDEG--------LYEAPKG-YTLEDLLATVK--YYVEHPNER-----
tr|Q9ANY3|Q9ANY3_STRPN    FEWFDEG--------LYEAPKG-YTLEDLLATVK--YYVEHPNER-----
                          *   *.          : .*:* :*  ::  .:    *:*:  :


tr|Q8DQ07|Q8DQ07_STRR6    PIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVEVP----ILEKENQT
tr|Q6WNQ7|Q6WNQ7_STRPN    PIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVEVP----ILEKENQT
tr|Q9ANY1|Q9ANY1_STRPN    PIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVEVP----ILEKENQT
tr|Q6WNQ5|Q6WNQ5_STRPN    PIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVEVP----ILEKENQT
tr|Q8DPQ2|Q8DPQ2_STRR6    --PHSNDG----WGN-------ASEHVLGKK---------------DHS
tr|Q9AG74|Q9AG74_STRPN    --PHSDNG----FGN-------ASDHVQRNKNGQ---------ADTNQT
tr|Q9AHT9|Q9AHT9_STRPN    --PHSNDG----WGN-------ASEHVLGKK---------------DHS
tr|Q8DQ08|Q8DQ08_STRR6    --PHSDNG----FGN-------ASDHVQRNKNGQADTNQTEKPNEEKPQT
tr|Q9ANY2|Q9ANY2_STRPN    --PHSDNG----FGN-------ASDHVRKNK--------------VDQD
tr|Q9ANY3|Q9ANY3_STRPN    --PHSDNG----FGN-------ASDHVQRNKNGQ---------ADTNQT
                          *: ::*       **       *. ::  :.                 :


tr|Q8DQ07|Q8DQ07_STRR6    DKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTL
tr|Q6WNQ7|Q6WNQ7_STRPN    DKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTL
tr|Q9ANY1|Q9ANY1_STRPN    DKPSILPQFKRNKAQENLKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTL
tr|Q6WNQ5|Q6WNQ5_STRPN    DKPSILPQFKRNKAQENSKFDEKVEEPKTSEKVEKEKLSETGNSTSNSTL
tr|Q8DPQ2|Q8DPQ2_STRR6    EDP----------NKNFKADEEPVE--ETP-AEP--------------
tr|Q9AG74|Q9AG74_STRPN    EKP----------NEEKPQTEKPEE--ETPREEKP-QSEKPESP-----
tr|Q9AHT9|Q9AHT9_STRPN    EDP----------NKNFKADEEPVE--ETP-AEP--------------
tr|Q8DQ08|Q8DQ08_STRR6    EKP----------EEDKEHDEVSEP--THPESDEK-ENHVGLNPS-ADN
tr|Q9ANY2|Q9ANY2_STRPN    SKP----------DEDKEHDEVSEP--THPESDEK-ENHAGLNPS-ADN
tr|Q9ANY3|Q9ANY3_STRPN    EKP----------SEEKPQTEKPEE--ETPREEKP-QSEKPESP-----
                          ..*            .::        *               :


tr|Q8DQ07|Q8DQ07_STRR6    EEVPTVDPVQEKVAKFAESYG-MKLENVLFNMDGTIELYLPSGEVIKKNM
tr|Q6WNQ7|Q6WNQ7_STRPN    EEVPTVDPVQEKVAKFAESYG-MKLENVLFNMDGTIELYLPSGEVIKKNM
tr|Q9ANY1|Q9ANY1_STRPN    EEVPTVDPVQEKVAKFAESYG-MKLENVLFNMDGTIELYLPSGEVIKKNM
tr|Q6WNQ5|Q6WNQ5_STRPN    EEVPTVDPVQEKVAKFAESYG-MKLENVLFNMDGTIELYLPSGEVIKKNM
tr|Q8DPQ2|Q8DPQ2_STRR6    -----------------E------VPQVET---EKVEAQLKEAEVLLAKV
tr|Q9AG74|Q9AG74_STRPN    --KPTEEPEEESPEESPE---ESEEPQVET---EKVKEKLREAEDLLGKI
tr|Q9AHT9|Q9AHT9_STRPN    -----------------E------VPQVET---EKVEAQLKEAEVLLAKV
tr|Q8DQ08|Q8DQ08_STRR6    LYKPSTDTEE-TEEEAEDTTDEAEIPQVEH---SVINAKIAEAEALLEKV
tr|Q9ANY2|Q9ANY2_STRPN    LYKPSTDTEE-TEEEAEDTTDEAEIPQVEN---SVINAKIADAEALLEKV
tr|Q9ANY3|Q9ANY3_STRPN    --KPTEEPEE-SPEESEE-------PQVET---EKVEEKLREAEDLLGKI
                            :             :*         :: :  ..* :  ::


tr|Q8DQ07|Q8DQ07_STRR6    ADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVK
tr|Q6WNQ7|Q6WNQ7_STRPN    ADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVK
tr|Q9ANY1|Q9ANY1_STRPN    ADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVK
tr|Q6WNQ5|Q6WNQ5_STRPN    ADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVK
tr|Q8DPQ2|Q8DPQ2_STRR6    TDSS--LKANATETLAGLRNNLTLQIMDNN-SIMAEAEKLLALLKGS--N
tr|Q9AG74|Q9AG74_STRPN    QNPI--IKSNAKETLTGLKNNLLFGTQDNN-TIMAEAEKLLALLKES--K
tr|Q9AHT9|Q9AHT9_STRPN    TDSS--LKANATETLAGLRNNLTLQIMDNN-SIMAEAEKLLALLKGS--N
tr|Q8DQ08|Q8DQ08_STRR6    TDSS--IRQNAVETLTGLKSSLLLGTKDNN-TISAEVDSLLALLKES--Q
tr|Q9ANY2|Q9ANY2_STRPN    TDPS--IRQNAMETLTGLKSSLLLGTKDNN-TISAEVDSLLALLKES--Q
```

```
tr|Q9ANY3|Q9ANY3_STRPN     QDPI--IKSNAKETLTGLKNNLLFGTQDNN-TIMAEAEKLLALLKES--K
                           :        . *. .. ...:     :*: :    .:.*    :  . :


tr|Q8DQ07|Q8DQ07_STRR6     PENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGL
tr|Q6WNQ7|Q6WNQ7_STRPN     PENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGL
tr|Q9ANY1|Q9ANY1_STRPN     PENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGL
tr|Q6WNQ5|Q6WNQ5_STRPN     PENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGL
tr|Q8DPQ2|Q8DPQ2_STRR6     PSSVSKEKIN----------------------------------------
tr|Q9AG74|Q9AG74_STRPN     --------------------------------------------------
tr|Q9AHT9|Q9AHT9_STRPN     PSSVSKEKIN----------------------------------------
tr|Q8DQ08|Q8DQ08_STRR6     PTPIQ---------------------------------------------
tr|Q9ANY2|Q9ANY2_STRPN     PAPIQ---------------------------------------------
tr|Q9ANY3|Q9ANY3_STRPN     --------------------------------------------------


tr|Q8DQ07|Q8DQ07_STRR6     DSVIFNMDGTIELRLPSGEVIKKNLSDLIA
tr|Q6WNQ7|Q6WNQ7_STRPN     DSVIFNMDGTIELRLPSGEVIKKNLSDLIA
tr|Q9ANY1|Q9ANY1_STRPN     DSVIFNMDGTIELRLPSGEVIKKNLSDLIA
tr|Q6WNQ5|Q6WNQ5_STRPN     DSVIFNMDGTIELRLPSGEVIKKNLSDLIA
tr|Q8DPQ2|Q8DPQ2_STRR6     ------------------------------
tr|Q9AG74|Q9AG74_STRPN     ------------------------------
tr|Q9AHT9|Q9AHT9_STRPN     ------------------------------
tr|Q8DQ08|Q8DQ08_STRR6     ------------------------------
tr|Q9ANY2|Q9ANY2_STRPN     ------------------------------
tr|Q9ANY3|Q9ANY3_STRPN     ------------------------------
```

PileUp

     MSF: 1080 Type: P   Check: 8540   ..

 Name: tr|Q8DQ07|Q8DQ07_STRR6 oo Len: 1080 Check: 6992 Weight: 0.100
 Name: tr|Q6WNQ7|Q6WNQ7_STRPN oo Len: 1080 Check: 6992 Weight: 0.100
 Name: tr|Q9ANY1|Q9ANY1_STRPN oo Len: 1080 Check: 7347 Weight: 0.100
 Name: tr|Q6WNQ5|Q6WNQ5_STRPN oo Len: 1080 Check: 4063 Weight: 0.100
 Name: tr|Q8DPQ2|Q8DPQ2_STRR6 oo Len: 1080 Check: 1836 Weight: 0.100
 Name: tr|Q9AG74|Q9AG74_STRPN oo Len: 1080 Check: 8409 Weight: 0.100
 Name: tr|Q9AHT9|Q9AHT9_STRPN oo Len: 1080 Check: 9461 Weight: 0.100
 Name: tr|Q8DQ08|Q8DQ08_STRR6 oo Len: 1080 Check: 9939 Weight: 0.100
 Name: tr|Q9ANY2|Q9ANY2_STRPN oo Len: 1080 Check: 1016 Weight: 0.100
 Name: tr|Q9ANY3|Q9ANY3_STRPN oo Len: 1080 Check: 2485 Weight: 0.100

//

tr|Q8DQ07|Q8DQ07_STRR6   ....MKFSKKYI AAGSAVIVSL SLCAYALNQH RSQENK.DNN
tr|Q6WNQ7|Q6WNQ7_STRPN   ....MKFSKKYI AAGSAVIVSL SLCAYALNQH RSQENK.DNN
tr|Q9ANY1|Q9ANY1_STRPN   ....MKFSKKYI AAGSAVIVSL SLCAYALNQH RSQENK.DNN
tr|Q6WNQ5|Q6WNQ5_STRPN   ....CAYALNQH RSQENK.DNN
tr|Q8DPQ2|Q8DPQ2_STRR6   MQLEISNRKR VSMKINKKYL VG.SAAALIL SVCSYELGLY QARTVK.ENN
tr|Q9AG74|Q9AG74_STRPN   ....MKINKKYL VG.SAAALIL SVCSYELGLY QARTVK.ENN
tr|Q9AHT9|Q9AHT9_STRPN   ....MKINKKYL VG.SAAALIL SVCSYELGLY QARTVK.ENN
tr|Q8DQ08|Q8DQ08_STRR6   ....MKINKKYL AG.SVAVLAL SVCSYELGRH QAGQVKKESN
tr|Q9ANY2|Q9ANY2_STRPN   ....MKINKKYL AG.SVAVLAL SVCSYELGRH QAGQVKKESN
tr|Q9ANY3|Q9ANY3_STRPN   ....MKINKKYL AG.SVAVLAL SVCSYELGRY QAGQDKKESN

tr|Q8DQ07|Q8DQ07_STRR6   RVSYVDGSQS SQKSENLTPD QVSQKEGIQA EQIVIKITDQ GYVTSHGDHY
tr|Q6WNQ7|Q6WNQ7_STRPN   RVSYVDGSQS SQKSENLTPD QVSQKEGIQA EQIVIKITDQ GYVTSHGDHY
tr|Q9ANY1|Q9ANY1_STRPN   RVSYVDGSQS SQKSENLTPD QVSQKEGIQA EQIVIKITDQ GYVTSHGDHY
tr|Q6WNQ5|Q6WNQ5_STRPN   RVSYVDGSQS SQKSENLTPD QVSQKEGIQA EQIVIKITDQ GYVTSHGDHY
tr|Q8DPQ2|Q8DPQ2_STRR6   RVSYIDGKQA TQKTENLTPD EVSKREGINA EQIVIKITDQ GYVTSHGDHY
tr|Q9AG74|Q9AG74_STRPN   RVSYIDGKQA TQKTENLTPD EVSKREGINA EQIVIKITDQ GYVTSHGDHY
tr|Q9AHT9|Q9AHT9_STRPN   RVSYIDGKQA TQKTENLTPD EVSKREGINA EQIVIKITDQ GYVTSHGDHY
tr|Q8DQ08|Q8DQ08_STRR6   RVSYIDGDQA GQKAENLTPD EVSKREGINA EQIVIKITDQ GYVTSHGDHY
tr|Q9ANY2|Q9ANY2_STRPN   RVSYIDGDQA GQKAENLTPD EVSKREGINA EQIVIKITDQ GYVTSHGDHY
tr|Q9ANY3|Q9ANY3_STRPN   RVAYIDGDQA GQKAENLTPD EVSKREGINA EQIVIKITDQ GYVTSHGDHY

tr|Q8DQ07|Q8DQ07_STRR6   HYYNGKVPYD ALFSEELLMK DPNYQLKDAD IVNEVKGGYI IKVDGKYYVY
tr|Q6WNQ7|Q6WNQ7_STRPN   HYYNGKVPYD ALFSEELLMK DPNYQLKDAD IVNEVKGGYI IKVDGKYYVY
tr|Q9ANY1|Q9ANY1_STRPN   HYYNGKVPYD ALFSEELLMK DPNYQLKDAD IVNEVKGGYI IKVDGKYYVY
tr|Q6WNQ5|Q6WNQ5_STRPN   HYYNGKVPYD ALFSEELLMK DPNYQLKDAD IVNEVKGGYI IKVDGKYYVY
tr|Q8DPQ2|Q8DPQ2_STRR6   HYYNGKVPYD AIFSEELLMK DPNYKLKDED IVNEVKGGYV IKVDGKYYVY
tr|Q9AG74|Q9AG74_STRPN   HYYNGKVPYD AIISEELLMK DPNYQLKDED IISEIKGGYV IKVDGKYYVY
tr|Q9AHT9|Q9AHT9_STRPN   HYYNGKVPYD AIISEELLMK DPNYKLKDED IVNEVKGGYV IKVDGKYYVY
tr|Q8DQ08|Q8DQ08_STRR6   HYYNGKVPYD AIISEELLMK DPNYQLKDSD IVNEIKGGYV IKVDGKYYVY
tr|Q9ANY2|Q9ANY2_STRPN   HYYNGKVPYD AIISEELLMK DPNYQLKDSD IVNEIKGGYV IKVDGKYYVY
tr|Q9ANY3|Q9ANY3_STRPN   HYYNGKVPYD AIISEELLMK DPNYQLKDSD IVNEIKGGYV IKVNGKYYVY

tr|Q8DQ07|Q8DQ07_STRR6   LKDAAHADNV RTKDEINRQK QEHVKDNE.. ..KVNSNVAV ARSQGRYTTN
tr|Q6WNQ7|Q6WNQ7_STRPN   LKDAAHADNV RTKDEINRQK QEHVKDNE.. ..KVNSNVAV ARSQGRYTTN
tr|Q9ANY1|Q9ANY1_STRPN   LKDAAHADNV RTKDEINRQK QEHVKDNE.. ..KVNSNVAV ARSQGRYTTN

```
tr|Q6WNQ5|Q6WNQ5_STRPN    LKDAAHADNV RTKDEINRQK QEHVKDNE.. ..KVNSNVAV ARSQGRYTTN
tr|Q8DPQ2|Q8DPQ2_STRR6    LKDAAHADNV RTKEEINRQK QEHSQHREGG TPRNDGAVAL ARSQGRYTTD
tr|Q9AG74|Q9AG74_STRPN    LKDAAHADNV RTKEEINRQK QEHSQHREGG TPRNDGAVAL ARSQGRYTTD
tr|Q9AHT9|Q9AHT9_STRPN    LKDAAHADNV RTKEEINRQK QEHSQHREGG TPRNDGAVAL ARSQGRYTTD
tr|Q8DQ08|Q8DQ08_STRR6    LKDAAHADNI RTKEEIKRQK QERSHNHN.. .SRADNAVAA ARAQGRYTTD
tr|Q9ANY2|Q9ANY2_STRPN    LKDAAHADNI RTKEEIKRQK QEHSHNHGG. .GSNDQAVVA ARAQGRYTTD
tr|Q9ANY3|Q9ANY3_STRPN    LKDAAHADNI RTKEEIKRQK QERSHNHN.. .SRADNAVAA ARAQGRYTTD


tr|Q8DQ07|Q8DQ07_STRR6    DGYVFNPADI IEDTGNAYIV PHGGHYHYIP KSDLSASELA AAKAHLAGK.
tr|Q6WNQ7|Q6WNQ7_STRPN    DGYVFNPADI IEDTGNAYIV PHGGHYHYIP KSDLSASELA AAKAHLAGK.
tr|Q9ANY1|Q9ANY1_STRPN    DGYVFNPADI IEDTGNAYIV PHGGHYHYIP KSDLSASELA AAKAHLAGK.
tr|Q6WNQ5|Q6WNQ5_STRPN    DGYVFNPADI IEDTGNAYIV PHRGHYHYIP KSDLSASELA AAKAHLAGK.
tr|Q8DPQ2|Q8DPQ2_STRR6    DGYIFNASDI IEDTGDAYIV PHGDHYHYIP KNELSASELA AAEAFLSGRG
tr|Q9AG74|Q9AG74_STRPN    DGYIFNASDI IEDTGDAYIV PHGDHYHYIP KNELSASELA AAKAFLSGRG
tr|Q9AHT9|Q9AHT9_STRPN    DGYIFNASDI IEDTGDAYIV PHGDHYHYIP KNELSASELA AAEAFLSGRG
tr|Q8DQ08|Q8DQ08_STRR6    DGYIFNASDI IEDTGDAYIV PHGDHYHYIP KSDLSASELA AAQAYWNGK.
tr|Q9ANY2|Q9ANY2_STRPN    DGYIFNASDI IEDTGDAYIV PHGDHYHYIP KNELSASELA AAEAYWNGK.
tr|Q9ANY3|Q9ANY3_STRPN    DGYIFNASDI IEDTGDAYIV PHGDHYHYIP KNELSASELA AAEAYWNGK.


tr|Q8DQ07|Q8DQ07_STRR6    .......... .NMQP:SQLS YSSTASD... NNTQSVAKGS TSKPANKSEN
tr|Q6WNQ7|Q6WNQ7_STRPN    .......... .NMQP.SQLS YSSTASD... NNTQSVAKGS TSKPANKSEN
tr|Q9ANY1|Q9ANY1_STRPN    .......... .NMQP.SQLS YSSTASD... NNTQSVAKGS TSKPANKSEN
tr|Q6WNQ5|Q6WNQ5_STRPN    .......... .NMQP.SQLS YSSTASD... NNTQSVAKGS TSKPANKSEN
tr|Q8DPQ2|Q8DPQ2_STRR6    NLSNSRTYRR QNSDNTSRTN WVPSVSNPGT TNTNTSNNSN TNSQASQSND
tr|Q9AG74|Q9AG74_STRPN    NLSNSRTYRR QNSDNTSRTN WVPSVSNPGT TNTNTSNNSN TNSQASQSND
tr|Q9AHT9|Q9AHT9_STRPN    NLSNSRTYRR QNSDNTSRTN WVPSVSNPGT TNTNTSNNSN TNSQASQSND
tr|Q8DQ08|Q8DQ08_STRR6    .......... QGSRPSSSSS HNANPAQPRL SENHNLTVTP TYHQN.QGEN
tr|Q9ANY2|Q9ANY2_STRPN    .......... QGSRPSSSSS YNANPAQPRL SENHNLTVTP TYHQN.QGEN
tr|Q9ANY3|Q9ANY3_STRPN    .......... QGSRPSSSSS YNANPAQPRL SENHNLTVTP TYHQN.QGEN


tr|Q8DQ07|Q8DQ07_STRR6    LQSLLKELYD SPSAQRYSES DGLVFDPAKI ISRTPNGVAI PHGDHYHFIP
tr|Q6WNQ7|Q6WNQ7_STRPN    LQSLLKELYD SPSAQRYSES DGLVFDPAKI ISRTPNGVAI PHGDHYHFIP
tr|Q9ANY1|Q9ANY1_STRPN    LQSLLKELYD SPSAQRYSES DGLVFDPAKI ISRTPNGVAI PHGDHYHFIP
tr|Q6WNQ5|Q6WNQ5_STRPN    LQSLLKELYD SPSAQRYSES DGLVFDPAKI ISRTPNGVAI PHGDHYHFIP
tr|Q8DPQ2|Q8DPQ2_STRR6    IDSLLKQLYK LPLSQRHVES DGLVFDPAQI TSRTARGVAV PHGDHYHFIP
tr|Q9AG74|Q9AG74_STRPN    IDSLLKQLYK LPLSQRHVES DGLIFDPAQI TSRTANGVAV PHGDHYHFIP
tr|Q9AHT9|Q9AHT9_STRPN    IDSLLKQLYK LPLSQRHVES DGLVFDPAQI TSRTARGVAV PHGDHYHFIP
tr|Q8DQ08|Q8DQ08_STRR6    ISSLLRELYA KPLSERHVES DGLIFDPAQI TSRTANGVAV PHGDHYHFIP
tr|Q9ANY2|Q9ANY2_STRPN    ISSLLRELYA KPLSERHVES DGLIFDPAQI TSRTARGVAV PHGNHYHFIP
tr|Q9ANY3|Q9ANY3_STRPN    ISSLLRELYA KPLSERHVES DGLIFDPAQI TSRTARGVAV PHGNHYHFIP


tr|Q8DQ07|Q8DQ07_STRR6    YSKLSALEEK IARRVPISGT GSTVSTNAKP NEVVSSLGSL SSNPS...SL
tr|Q6WNQ7|Q6WNQ7_STRPN    YSKLSALEEK IARRVPISGT GSTVSTNAKP NEVVSSLGSL SSNPS...SL
tr|Q9ANY1|Q9ANY1_STRPN    YSKLSALEEK IARMVPISGT GSTVSTNAKP NEVVSSLGSL SSNPS...SL
tr|Q6WNQ5|Q6WNQ5_STRPN    YSKLSALEEK IARMVPISGT GSTVSTNAKP NEVVSSLGSL SSNPS...SL
tr|Q8DPQ2|Q8DPQ2_STRR6    YSQMSELEER IARIIPLRYR SNHWVPDSRP .EQPSPQPTP EPSPG.....
tr|Q9AG74|Q9AG74_STRPN    YSQLSPLEEK LARIIPLRYR SNHWVPDSRP .EQPSPQSTP EPSPSPQPAP
tr|Q9AHT9|Q9AHT9_STRPN    YSQMSELEER IARIIPLRYR SNHWVPDSRP .EQPSPQPTP EPSPG.....
tr|Q8DQ08|Q8DQ08_STRR6    YSQLSPLEEK LARIIPLRYR SNHWVPDSRP .EQPSPQSTP EPSPSPQPAP
tr|Q9ANY2|Q9ANY2_STRPN    YEQMSELEKR IARIIPLRYR SNHWVPDSRP .EQPSPQSTP EPSPSPQPAP
tr|Q9ANY3|Q9ANY3_STRPN    YEQMSELEKR IARIIPLRYR SNHWVPDSRP .EEPSPQPTP EPSPS.....


tr|Q8DQ07|Q8DQ07_STRR6    TTSKELSSAS DGYIFNPKDI VEETATAYIV RHGDHFHYIP KSNQIGQPTL
tr|Q6WNQ7|Q6WNQ7_STRPN    TTSKELSSAS DGYIFNPKDI VEETATAYIV RHGDHFHYIP KSNQIGQPTL
tr|Q9ANY1|Q9ANY1_STRPN    TTSKELSSAS DGYIFNPKDI VEETATAYIV RHGDHFHYIP KSNQIGQPTL
```

```
tr|Q6WNQ5|Q6WNQ5_STRPN    TTSKELSSAS DGYIFNPKDI VEETATAYIV RHGDHFHYIP KSNQIGQPTL
tr|Q8DPQ2|Q8DPQ2_STRR6    .PQPAPNLKI DSNSSLVSQL VRKVGEGYVF EEKGISRYVF AKD......L
tr|Q9AG74|Q9AG74_STRPN    NPQPAPSNPI DEK..LVKEA VRKVGDGYVF EENGVPRYIP AKD......L
tr|Q9AHT9|Q9AHT9_STRPN    .PQPAPNLKI DSNSSLVSQL VRKVGEGYVF EEKGISRYVF AKD......L
tr|Q8DQ08|Q8DQ08_STRR6    NPQPAPSNPI DEK..LVKEA VRKVGDGYVF EENGVPRYIP AKD......L
tr|Q9ANY2|Q9ANY2_STRPN    NPQPAPSNPI DEK..LVKEA VRKVGDGYVF EENGVSRYIP AKD......L
tr|Q9ANY3|Q9ANY3_STRPN    .PQPAPSNPI DGK..LVKEA VRKVGDGYVF EENGVSRYIP AKD......L


tr|Q8DQ07|Q8DQ07_STRR6    PNNSLATPSP SLPINPGISH EKHEEDGYGF DANRIIAEDE SGFIMSHGNH
tr|Q6WNQ7|Q6WNQ7_STRPN    PNNSLATPSP SLPINPGISH EKHEEDGYGF DANRIIAEDE SGFIMSHGNH
tr|Q9ANY1|Q9ANY1_STRPN    PNNSLATPSP SLPINPGTSH EKHEEDGYGF DANRIIAEDE SGFVMSHGDH
tr|Q6WNQ5|Q6WNQ5_STRPN    PNNSLATPSP SLPINPGTSH EKHEEDGYGF DANRIIAEDE SGFVMSHGDH
tr|Q8DPQ2|Q8DPQ2_STRR6    PSETVKNLES KLSKQESVSH T......... ....LTAKKE N...VAPRDQ
tr|Q9AG74|Q9AG74_STRPN    SAETAAGIDS KLAKQESLSH K......... ....LGAKKT D...LPSSDR
tr|Q9AHT9|Q9AHT9_STRPN    PSETVKNLES KLSKQESVSH T......... ....LTAKKE ·N...VAPRDQ
tr|Q8DQ08|Q8DQ08_STRR6    SAETAAGIDS KLAKQESLSH K......... ....LGAKKT D...LPSSDR
tr|Q9ANY2|Q9ANY2_STRPN    SAETAAGIDS KLAKQESLSH K......... ....LGAKKT D...LPSSDR
tr|Q9ANY3|Q9ANY3_STRPN    SAETAAGIDS KLAKQESLSH K......... ....LGTKKT D...LPSSDR


tr|Q8DQ07|Q8DQ07_STRR6    NHYFFKKDLT EEQIKAAQKH LEEVKTSHNG LDSLSSHEQD YPGNAKEMKD
tr|Q6WNQ7|Q6WNQ7_STRPN    NHYFFKKDLT EEQIKAAQKH LEEVKTSHNG LDSLSSHEQD YPGNAKEMKD
tr|Q9ANY1|Q9ANY1_STRPN    NHYFFKKDLT EEQIKAAQKH LEEVKTSHNG LDSLSSHEQD YPSNAKEMKD
tr|Q6WNQ5|Q6WNQ5_STRPN    NHYFFKKDLT EEQIKAAQKH LEEVKTSHNG LDSLSSHEQD YPSNAKEMKD
tr|Q8DPQ2|Q8DPQ2_STRR6    EFYDKAYNLL TEAHKALFEN .KGRNSDFQA LDKLLERLND EST......N
tr|Q9AG74|Q9AG74_STRPN    EFYNKAYDLL ARIHQDLLDN .KGRQVDFEA LDNLLERLKD VSS......D
tr|Q9AHT9|Q9AHT9_STRPN    EFYDKAYNLL TEAHKALFXN .KGRNSDFQA LDKLLERLND EST......N
tr|Q8DQ08|Q8DQ08_STRR6    EFYNKAYDLL ARIHQDLLDN .KGRQVDFEA LDNLLERLKD VSS......D
tr|Q9ANY2|Q9ANY2_STRPN    EFYNKAYDLL ARIHQDLLDN .KGRQVDFEA LDNLLERLKD VPS......D
tr|Q9ANY3|Q9ANY3_STRPN    EFYNKAYDLL ARIHQDLLDN .KGRQVDFEA LDNLLERLKD VSS......D


tr|Q8DQ07|Q8DQ07_STRR6    LDKKIEEKIA GIMKQYGVKR ESIVVNKEKN AIIYPHGDHH HADPIDEHKP
tr|Q6WNQ7|Q6WNQ7_STRPN    LDKKIEEKIA GIMKQYGVKR ESIVVNKEKN AIIYPHGDHH HADPIDEHKP
tr|Q9ANY1|Q9ANY1_STRPN    LDKKIEEKIA GIMKQYGVKR ESIVVNKEKN AIIYPHGDHH HADPIDEHKP
tr|Q6WNQ5|Q6WNQ5_STRPN    LDKKIEEKIA GIMKQYGVKR ESIVVNKEKN AIIYPHGDHH HADPIDEHKP
tr|Q8DPQ2|Q8DPQ2_STRR6    KEKLVDDLLA FLAPITHPER ....LGKPNS QIEYTE.... ..DEVRIAQL
tr|Q9AG74|Q9AG74_STRPN    KVKLVDDILA FLAPIRHPER ....LGKPNA QITYTD.... ..DEIQVAKL
tr|Q9AHT9|Q9AHT9_STRPN    KEKLVDDLLA FLAPITHPER ....LGKPNS QIEYTE.... ..DEVRIAQL
tr|Q8DQ08|Q8DQ08_STRR6    KVKLVDDILA FLAPIRHPER ....LGKPNA QITYTD.... ..DEIQVAKL
tr|Q9ANY2|Q9ANY2_STRPN    KVKLVDDILA FLAPIRHPER ....LGKPNA QITYTD.... ..DEIQVAKL
tr|Q9ANY3|Q9ANY3_STRPN    KVKLVEDILA FLAPIRHPER ....LGKPNA QITYTD.... ..DEIQVAKL


tr|Q8DQ07|Q8DQ07_STRR6    VGIGHSHSNY ELFKPEEGVA KKEGNKVYTG EELTNVVNLL KNSTFNNQNF
tr|Q6WNQ7|Q6WNQ7_STRPN    VGIGHSHSNY ELFKPEEGVA KKEGNKVYTG EELTNVVNLL KNSTFNNQNF
tr|Q9ANY1|Q9ANY1_STRPN    VGIGHSHSNY ELFKPEEGVA KKEGNKVYTG EELTNVVNLL KNSTFNNQNF
tr|Q6WNQ5|Q6WNQ5_STRPN    VGIGHSHSNY ELFKPEEGVA KKEGNKVYTG EELTNVVNLL KNSTFNNQNF
tr|Q8DPQ2|Q8DPQ2_STRR6    ADK.YTTSDG YIFDEHD.II SDEGD.AYVT PHMGHS.HWI GKDSLSDKEK
tr|Q9AG74|Q9AG74_STRPN    AGK.YTTEDG YIFDPRD.IT SDEGD.AYVT PHMTHS.HWI KKDSLSEAER
tr|Q9AHT9|Q9AHT9_STRPN    ADK.YTTSDG YIFDEHD.II SDEGD.AYVT PHMGHS.HWI GKDSLSDKEK
tr|Q8DQ08|Q8DQ08_STRR6    AGK.YTTEDG YIFDPRD.IT SDEGD.AYVT PHMTHS.HWI KKDSLSEAER
tr|Q9ANY2|Q9ANY2_STRPN    AGK.YTTEDG YIFDPRD.IT SDEGD.AYVT PHMTHS.HWI KKDSLSEAER
tr|Q9ANY3|Q9ANY3_STRPN    AGK.YTAEDG YIFDPRD.IT SDEGD.AYVT PHMTHS.HWI KKDSLSEAER


tr|Q8DQ07|Q8DQ07_STRR6    TLANGQKRVS FSFPPELEKK LGINMLVKLI TPDGKVLEKV SGKVFGEGVG
tr|Q6WNQ7|Q6WNQ7_STRPN    TLANGQKRVS FSFPPELEKK LGINMLVKLI TPDGKVLEKV SGKVFGEGVG
tr|Q9ANY1|Q9ANY1_STRPN    TLANGQKRVS FSFPPELEKK LGINMLVKLI TPDGKVLEKV SGKVFGEGVG
```

```
tr|Q6WNQ5|Q6WNQ5_STRPN    TLANGQKRVS FSFPPELEKK LGINMLVKLI TPDGKVLEKV SGKVFGEGVG
tr|Q8DPQ2|Q8DPQ2_STRR6    VAAQAYTKEK GILPPSPDAD VKAN...... .PTGDSAAAI YNRVKGE...
tr|Q9AG74|Q9AG74_STRPN    AAAQAYAKEK GLTPPSTDHQ DSGN...... .TEAKGAEAI YNRVKAA...
tr|Q9AHT9|Q9AHT9_STRPN    VAAQAYTKEK GILPPSPDAD VKAN...... .PTGDSAAAI YNRVKGE...
tr|Q8DQ08|Q8DQ08_STRR6    AAAQAYAKEK GLTPPSTDHQ DSGN...... .TEAKGAEAI YNRVKAA...
tr|Q9ANY2|Q9ANY2_STRPN    AAAQAYAKEK GLTPPSTDHQ DSGN...... .TEAKGAEAI YNRVKAA...
tr|Q9ANY3|Q9ANY3_STRPN    AAAQAYAXEK GLTPPSTDHQ DSGN...... .TEAKGAEAI YNRVKAA...


tr|Q8DQ07|Q8DQ07_STRR6    NIANFELDQP YLPGQTFKYT IASKDYPEVS YDGTFTVPTS LAYKMASQTI
tr|Q6WNQ7|Q6WNQ7_STRPN    NIANFELDQP YLPGQTFKYT IASKDYPEVS YDGTFTVPTS LAYKMASQTI
tr|Q9ANY1|Q9ANY1_STRPN    NIANFELDQP YLPGQTFKYT IASKDYPEVS YDGTFTVPTS LAYKMASQTI
tr|Q6WNQ5|Q6WNQ5_STRPN    NIANFELDQP YLPGQTFKYT IASKDYPEVS YDGTFTVPTS LAYKMASQTI
tr|Q8DPQ2|Q8DPQ2_STRR6    ..KRIPLVR. .LP.YMVEHT VEVK...... .NGNLIIP.. .HKDHYHNIK
tr|Q9AG74|Q9AG74_STRPN    ..KKVPLDR. .MP.YNLQYT VEVK...... .NGSLIIP.. .HYDHYHNIK
tr|Q9AHT9|Q9AHT9_STRPN    ..KRIPLVR. .LP.YMVEHT VEVK...... .NGNLIIP.. .HKDHYHNIK
tr|Q8DQ08|Q8DQ08_STRR6    ..KKVPLDR. .MP.YNLQYT VEVK...... .NGSLIIP.. .HYDHYHNIK
tr|Q9ANY2|Q9ANY2_STRPN    ..KKVPLDR. .MP.YNLQYT VEVK...... .NGSLIIP.. .HYDHYHNIK
tr|Q9ANY3|Q9ANY3_STRPN    ..KKVPLDR. .MP.YNLQYT VEVK...... .NGSLIIP.. .HYDHYHNIK


tr|Q8DQ07|Q8DQ07_STRR6    FYPFHAGDTY LRVNPQFAVP KGTDALVRVF DEFHGNAYLE NNYKVGEIKL
tr|Q6WNQ7|Q6WNQ7_STRPN    FYPFHAGDTY LRVNPQFAVP KGTDALVRVF DEFHGNAYLE NNYKVGEIKL
tr|Q9ANY1|Q9ANY1_STRPN    FYPFHAGDTY LRVNPQFAVP KGTDALVRVF DEFHGNAYLE NNYKVGEIKL
tr|Q6WNQ5|Q6WNQ5_STRPN    FYPFHAGDTY LRVNPQFAVP KGTDALVRVF DEFHGNAYLE NNYKVGEIKL
tr|Q8DPQ2|Q8DPQ2_STRR6    FAWFDDH... .....TYKAP NG.YTLEDLF ATIK..YYVE HPDER.....
tr|Q9AG74|Q9AG74_STRPN    FEWFDEG... .....LYEAP KG.YSLEDLL ATVK..YYVE HPNER.....
tr|Q9AHT9|Q9AHT9_STRPN    FAWFDDH... .....TYKAP NG.YTLEDLF ATIK..YYVE HPDER.....
tr|Q8DQ08|Q8DQ08_STRR6    FEWFDEG... .....LYEAP KG.YSLEDLL ATVK..YYVE HPNER.....
tr|Q9ANY2|Q9ANY2_STRPN    FEWFDEG... .....LYEAP KG.YTLEDLL ATVK..YYVE HPNER.....
tr|Q9ANY3|Q9ANY3_STRPN    FEWFDEG... .....LYEAP KG.YTLEDLL ATVK..YYVE HPNER.....


tr|Q8DQ07|Q8DQ07_STRR6    PIPKLNQGTT RTAGNKIPVT FMANAYLDNQ STYIVEVP.. ..ILEKENQT
tr|Q6WNQ7|Q6WNQ7_STRPN    PIPKLNQGTT RTAGNKIPVT FMANAYLDNQ STYIVEVP.. ..ILEKENQT
tr|Q9ANY1|Q9ANY1_STRPN    PIPKLNQGTT RTAGNKIPVT FMANAYLDNQ STYIVEVP.. ..ILEKENQT
tr|Q6WNQ5|Q6WNQ5_STRPN    PIPKLNQGTT RTAGNKIPVT FMANAYLDNQ STYIVEVP.. ..ILEKENQT
tr|Q8DPQ2|Q8DPQ2_STRR6    ..PHSNDG.. ..WGN..... ..ASEHVLGK K......... .......DHS
tr|Q9AG74|Q9AG74_STRPN    ..PHSDNG.. ..FGN..... ..ASDHVQRN KNGQ...... ....ADTNQT
tr|Q9AHT9|Q9AHT9_STRPN    ..PHSNDG.. ..WGN..... ..ASEHVLGK K......... .......DHS
tr|Q8DQ08|Q8DQ08_STRR6    ..PHSDNG.. ..FGN..... ..ASDHVQRN KNGQADTNQT EKPNEEKPQT
tr|Q9ANY2|Q9ANY2_STRPN    ..PHSDNG.. ..FGN..... ..ASDHVRKN K......... ......VDQD
tr|Q9ANY3|Q9ANY3_STRPN    ..PHSDNG.. ..FGN..... ..ASDHVQRN KNGQ...... ....ADTNQT


tr|Q8DQ07|Q8DQ07_STRR6    DKPSILPQFK RNKAQENSKL DEKVEEPKTS EKVEKEKLSE TGNSTSNSTL
tr|Q6WNQ7|Q6WNQ7_STRPN    DKPSILPQFK RNKAQENSKL DEKVEEPKTS EKVEKEKLSE TGNSTSNSTL
tr|Q9ANY1|Q9ANY1_STRPN    DKPSILPQFK RNKAQENLKL DEKVEEPKTS EKVEKEKLSE TGNSTSNSTL
tr|Q6WNQ5|Q6WNQ5_STRPN    DKPSILPQFK RNKAQENSKF DEKVEEPKTS EKVEKEKLSE TGNSTSNSTL
tr|Q8DPQ2|Q8DPQ2_STRR6    EDP....... ....NKNFKA DEEPVE..ET P.AEP..... ..........
tr|Q9AG74|Q9AG74_STRPN    EKP....... ....NEEKPQ TEKPEE..ET PREEKP.QSE KPESP.....
tr|Q9AHT9|Q9AHT9_STRPN    EDP....... ....NKNFKA DEEPVE..ET P.AEP..... ..........
tr|Q8DQ08|Q8DQ08_STRR6    EKP....... ....EEDKEH DEVSEP..TH PESDEK.ENH VGLNPS.ADN
tr|Q9ANY2|Q9ANY2_STRPN    SKP....... ....DEDKEH DEVSEP..TH PESDEK.ENH AGLNPS.ADN
tr|Q9ANY3|Q9ANY3_STRPN    EKP....... ....SEEKPQ TEKPEE..ET PREEKP.QSE KPESP.....


tr|Q8DQ07|Q8DQ07_STRR6    EEVPTVDPVQ EKVAKFAESY G.MKLENVLF NMDGTIELYL PSGEVIKKNM
tr|Q6WNQ7|Q6WNQ7_STRPN    EEVPTVDPVQ EKVAKFAESY G.MKLENVLF NMDGTIELYL PSGEVIKKNM
tr|Q9ANY1|Q9ANY1_STRPN    EEVPTVDPVQ EKVAKFAESY G.MKLENVLF NMDGTIELYL PSGEVIKKNM
```

```
tr|Q6WNQ5|Q6WNQ5_STRPN    EEVPTVDPVQ EKVAKFAESY G.MKLENVLF NMDGTIELYL PSGEVIKKNM
tr|Q8DPQ2|Q8DPQ2_STRR6    .......... .......E.. ....VPQVET ...EKVEAQL KEAEVLLAKV
tr|Q9AG74|Q9AG74_STRPN    ..KPTEEPEE ESPEESPE.. .ESEEPQVET ...EKVKEKL REAEDLLGKI
tr|Q9AHT9|Q9AHT9_STRPN    .......... .......E.. ....VPQVET ...EKVEAQL KEAEVLLAKV
tr|Q8DQ08|Q8DQ08_STRR6    LYKPSTDTEE .TEEEAEDTT DEAEIPQVEH ...SVINAKI AEAEALLEKV
tr|Q9ANY2|Q9ANY2_STRPN    LYKPSTDTEE .TEEEAEDTT DEAEIPQVEN ...SVINAKI ADAEALLEKV
tr|Q9ANY3|Q9ANY3_STRPN    ..KPTEEPEE .SPEESEE.. ......PQVET ...EKVEEKL REAEDLLGKI


tr|Q8DQ07|Q8DQ07_STRR6    ADFTGEAPQG NGENKPSENG KVSTGTVENQ PTENKPADSL PEAPNEKPVK
tr|Q6WNQ7|Q6WNQ7_STRPN    ADFTGEAPQG NGENKPSENG KVSTGTVENQ PTENKPADSL PEAPNEKPVK
tr|Q9ANY1|Q9ANY1_STRPN    ADFTGEAPQG NGENKPSENG KVSTGTVENQ PTENKPADSL PEAPNEKPVK
tr|Q6WNQ5|Q6WNQ5_STRPN    ADFTGEAPQG NGENKPSENG KVSTGTVENQ PTENKPADSL PEAPNEKPVK
tr|Q8DPQ2|Q8DPQ2_STRR6    TDSS..LKAN ATETLAGLRN NLTLQIMDNN .SIMAEAEKL LALLKGS..N
tr|Q9AG74|Q9AG74_STRPN    QNPI..IKSN AKETLTGLKN NLLFGTQDNN .TIMAEAEKL LALLKES..K
tr|Q9AHT9|Q9AHT9_STRPN    TDSS..LKAN ATETLAGLRN NLTLQIMDNN .SIMAEAEKL LALLKGS..N
tr|Q8DQ08|Q8DQ08_STRR6    TDSS..IRQN AVETLTGLKS SLLLGTKDNN .TISAEVDSL LALLKES..Q
tr|Q9ANY2|Q9ANY2_STRPN    TDPS..IRQN AMETLTGLKS SLLLGTKDNN .TISAEVDSL LALLKES..Q
tr|Q9ANY3|Q9ANY3_STRPN    QDPI..IKSN AKETLTGLKN NLLFGTQDNN .TIMAEAEKL LALLKES..K


tr|Q8DQ07|Q8DQ07_STRR6    PENSTDNGML NPEGNVGSDP MLDPALEEAP AVDPVQEKLE KFTASYGLGL
tr|Q6WNQ7|Q6WNQ7_STRPN    PENSTDNGML NPEGNVGSDP MLDPALEEAP AVDPVQEKLE KFTASYGLGL
tr|Q9ANY1|Q9ANY1_STRPN    PENSTDNGML NPEGNVGSDP MLDPALEEAP AVDPVQEKLE KFTASYGLGL
tr|Q6WNQ5|Q6WNQ5_STRPN    PENSTDNGML NPEGNVGSDP MLDPALEEAP AVDPVQEKLE KFTASYGLGL
tr|Q8DPQ2|Q8DPQ2_STRR6    PSSVSKEKIN .......... .......... .......... ..........
tr|Q9AG74|Q9AG74_STRPN    .......... .......... .......... .......... ..........
tr|Q9AHT9|Q9AHT9_STRPN    PSSVSKEKIN .......... .......... .......... ..........
tr|Q8DQ08|Q8DQ08_STRR6    PTPIQ..... .......... .......... .......... ..........
tr|Q9ANY2|Q9ANY2_STRPN    PAPIQ..... .......... .......... .......... ..........
tr|Q9ANY3|Q9ANY3_STRPN    .......... .......... .......... .......... ..........


tr|Q8DQ07|Q8DQ07_STRR6    DSVIFNMDGT IELRLPSGEV IKKNLSDLIA
tr|Q6WNQ7|Q6WNQ7_STRPN    DSVIFNMDGT IELRLPSGEV IKKNLSDLIA
tr|Q9ANY1|Q9ANY1_STRPN    DSVIFNMDGT IELRLPSGEV IKKNLSDLIA
tr|Q6WNQ5|Q6WNQ5_STRPN    DSVIFNMDGT IELRLPSGEV IKKNLSDLIA
tr|Q8DPQ2|Q8DPQ2_STRR6    .......... .......... ..........
tr|Q9AG74|Q9AG74_STRPN    .......... .......... ..........
tr|Q9AHT9|Q9AHT9_STRPN    .......... .......... ..........
tr|Q8DQ08|Q8DQ08_STRR6    .......... .......... ..........
tr|Q9ANY2|Q9ANY2_STRPN    .......... .......... ..........
tr|Q9ANY3|Q9ANY3_STRPN    .......... .......... ..........
```

Alignments

tr <u>Q8DQ07</u> **Pneumococcal histidine triad protein E [phtE] [Streptococcus 1039 AA
           pneumoniae (strain ATCC BAA-255 / R6)]**

<u>align</u>

  Score = 2017 bits (5225), Expect = 0.0
  Identities = 1004/1039 (96%), Positives = 1004/1039 (96%)

Query: 1   MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60
             MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS
Sbjct: 1   MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60

Query: 61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120
             QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN
Sbjct: 61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120

Query: 121   EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180
             EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY
Sbjct: 121   EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180

Query: 181   TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLS 240
             TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIP   NMQPSQLS
Sbjct: 181   TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLS 240

Query: 241   YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300
             YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
Sbjct: 241   YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300

Query: 301   TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXX 360
             TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVV
Sbjct: 301   TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVSSLGSLSSN 360

Query: 361   XXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420
                    KELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA
Sbjct: 361   PSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420

Query: 421   TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA 480
             TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA
Sbjct: 421   TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA 480

Query: 481   AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540
             AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN
Sbjct: 481   AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540

Query: 541   KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600
             KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
Sbjct: 541   KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600

Query: 601   VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660
             VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG
Sbjct: 601   VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660

Query: 661   EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720
             EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA
Sbjct: 661   EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720

```
Query: 721   GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780
             GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK
Sbjct: 721   GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780

Query: 781   IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS 840
             IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS
Sbjct: 781   IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS 840

Query: 841   EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900
             EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
Sbjct: 841   EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900

Query: 901   SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960
             SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP
Sbjct: 901   SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960

Query: 961   ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020
             ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI
Sbjct: 961   ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020

Query: 1021  ELRLPSGEVIKKNLSDLIA 1039
             ELRLPSGEVIKKNLSDLIA
Sbjct: 1021  ELRLPSGEVIKKNLSDLIA 1039
```

tr   Q6WNQ7   **Surface protein BVH-3 [bvh-3] [Streptococcus**   1039 AA
                    pneumoniae]   align

```
 Score = 2017 bits (5225), Expect = 0.0
 Identities = 1004/1039 (96%), Positives = 1004/1039 (96%)

Query: 1     MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60
             MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS
Sbjct: 1     MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60

Query: 61    QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120
             QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN
Sbjct: 61    QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120

Query: 121   EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180
             EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY
Sbjct: 121   EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180

Query: 181   TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQLS 240
             TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIP                     NMQPSQLS
Sbjct: 181   TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLS 240

Query: 241   YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300
             YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
Sbjct: 241   YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300

Query: 301   TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXX 360
             TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVV
Sbjct: 301   TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVSSLGSLSSN 360

Query: 361   XXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420
```

```
                  KELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA
Sbjct:  361   PSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA  420

Query:  421   TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA  480
              TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA
Sbjct:  421   TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA  480

Query:  481   AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN  540
              AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN
Sbjct:  481   AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN  540

Query:  541   KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV  600
              KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
Sbjct:  541   KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV  600

Query:  601   VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG  660
              VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG
Sbjct:  601   VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG  660

Query:  661   EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA  720
              EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA
Sbjct:  661   EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA  720

Query:  721   GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK  780
              GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK
Sbjct:  721   GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK  780

Query:  781   IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS  840
              IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS
Sbjct:  781   IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS  840

Query:  841   EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP  900
              EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
Sbjct:  841   EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP  900

Query:  901   SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP  960
              SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP
Sbjct:  901   SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP  960

Query:  961   ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI  1020
              ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI
Sbjct:  961   ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI  1020

Query: 1021   ELRLPSGEVIKKNLSDLIA  1039
              ELRLPSGEVIKKNLSDLIA
Sbjct: 1021   ELRLPSGEVIKKNLSDLIA  1039
```

```
tr  Q9ANY1   Pneumococcal histidine triad protein E precursor          1039
            (Hypothetical                                               AA
            protein SP1004) [phtE] [Streptococcus pneumoniae]          align
```

 Score = 2006 bits (5196), Expect = 0.0
 Identities = 998/1039 (96%), Positives = 1000/1039 (96%)

```
Query:  1     MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS  60
```

```
            MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS
Sbjct: 1   MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60

Query: 61  QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSBELLMKDPNYQLKDADIVN 120
            QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSBELLMKDPNYQLKDADIVN
Sbjct: 61  QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSBELLMKDPNYQLKDADIVN 120

Query: 121 EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180
            EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY
Sbjct: 121 EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180

Query: 181 TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQLS 240
            TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIP                     NMQPSQLS
Sbjct: 181 TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLS 240

Query: 241 YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300
            YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
Sbjct: 241 YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300

Query: 301 TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXX 360
            TPNGVAIPHGDHYHFIPYSKLSALEEKIAR VPISGTGSTVSTNAKPNEVV
Sbjct: 301 TPNGVAIPHGDHYHFIPYSKLSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSN 360

Query: 361 XXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420
                   KELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA
Sbjct: 361 PSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420

Query: 421 TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA 480
            TPSPSLPINPG SHEKHEEDGYGFDANRIIAEDESGF+MSHG+HNHYFFKKDLTEEQIKA
Sbjct: 421 TPSPSLPINPGTSHEKHEEDGYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKA 480

Query: 481 AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540
            AQKHLEEVKTSHNGLDSLSSHEQDYP NAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN
Sbjct: 481 AQKHLEEVKTSHNGLDSLSSHEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540

Query: 541 KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600
            KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
Sbjct: 541 KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600

Query: 601 VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660
            VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG
Sbjct: 601 VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660

Query: 661 EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720
            EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA
Sbjct: 661 EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720

Query: 721 GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780
            GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK
Sbjct: 721 GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780

Query: 781 IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS 840
            IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQEN KLDEKVEEPKTS
Sbjct: 781 IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENLKLDEKVEEPKTS 840

Query: 841 EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900
            EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
Sbjct: 841 EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900
```

```
Query:  901  SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960
             SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP
Sbjct:  901  SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960

Query:  961  ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020
             ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI
Sbjct:  961  ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020

Query: 1021  ELRLPSGEVIKKNLSDLIA 1039
             ELRLPSGEVIKKNLSDLIA
Sbjct: 1021  ELRLPSGEVIKKNLSDLIA 1039
```

tr <u>Q6WNQ5</u> **Surface protein BVH-3 (Fragment) [bvh-3] [Streptococcus pneumoniae]**

1019
AA
<u>align</u>

Score = 1968 bits (5099), Expect = 0.0
Identities = 977/1019 (95%), Positives = 979/1019 (95%)

```
Query:   21  CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV 80
             CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV
Sbjct:    1  CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV 60

Query:   81  TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD 140
             TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD
Sbjct:   61  TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD 120

Query:  141  AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA 200
             AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA
Sbjct:  121  AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA 180

Query:  201  YIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSK 260
             YIVPH GHYHYIP                       NMQPSQLSYSSTASDNNTQSVAKGSTSK
Sbjct:  181  YIVPHRGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSSTASDNNTQSVAKGSTSK 240

Query:  261  PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK 320
             PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK
Sbjct:  241  PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK 300

Query:  321  LSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXXXKELSSASDGYIFN 380
             LSALEEKIAR VPISGTGSTVSTNAKPNEVV                    KELSSASDGYIFN
Sbjct:  301  LSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSNPSSLTTSKELSSASDGYIFN 360

Query:  381  PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGISHEKHEED 440
             PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPG SHEKHEED
Sbjct:  361  PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGTSHEKHEED 420

Query:  441  GYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS 500
             GYGFDANRIIAEDESGF+MSHG+HNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS
Sbjct:  421  GYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS 480

Query:  501  HEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID 560
             HEQDYP NAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID
Sbjct:  481  HEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID 540
```

```
Query: 561  EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ 620
            EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ
Sbjct: 541  EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ 600

Query: 621  KRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT 680
            KRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT
Sbjct: 601  KRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT 660

Query: 681  FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL 740
            FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL
Sbjct: 661  FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL 720

Query: 741  VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE 800
            VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE
Sbjct: 721  VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE 780

Query: 801  VPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTL 860
            VPILEKENQTDKPSILPQFKRNKAQENSK DEKVEEPKTSEKVEKEKLSETGNSTSNSTL
Sbjct: 781  VPILEKENQTDKPSILPQFKRNKAQENSKFDEKVEEPKTSEKVEKEKLSETGNSTSNSTL 840

Query: 861  EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN 920
            EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN
Sbjct: 841  EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN 900

Query: 921  GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM 980
            GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM
Sbjct: 901  GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM 960

Query: 981  LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA 1039
            LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA
Sbjct: 961  LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA 1019
```

tr <u>Q8CWR4</u> **Histidine Motif-Containing protein [phpA] [Streptococcus**   855
              pneumoniae   AA
              (strain ATCC BAA-255 / R6)]   <u>align</u>

 Score = 442 bits (1137), Expect = e-122
 Identities = 219/369 (59%), Positives = 271/369 (73%), Gaps = 21/369 (5%)

```
Query: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60
            MK +KKY+A  A +V LS+CAY L H++Q  K+NNRVSY+DG Q++QK+ENLTPD+VS
Sbjct: 12   MKINKKYLAGSVATLV-LSVCAYELGLHQAQTVKENNRVSYIDGKQATQKTENLTPDEVS 70

Query: 61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120
            ++EGI AEQIVIKITDQGYVTSHGDHYHYYNGKVPYDA+ SEELLMKDPNYQLKD DI++
Sbjct: 71   KREGINAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDEDIIS 130

Query: 121  EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNE----KVNSNVAVARS 176
            E+KGGY+IKVDGKYYVYLKDAAHADNVRTK+EINRQKQEH + E  +  + VA+ARS
Sbjct: 131  EIKGGYVIKVDGKYYVYLKDAAHADNVRTKEEINRQKQEHSQHREGGTPRNDGAVALARS 190

Query: 177  QGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQP 236
            QGRYTT+DGY+FN +DIIEDTG+AYIVPHG HYHYIP
Sbjct: 191  QGRYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAKAFLSGRGNLS 250
```

Figure 5

A. Full-length Sp36
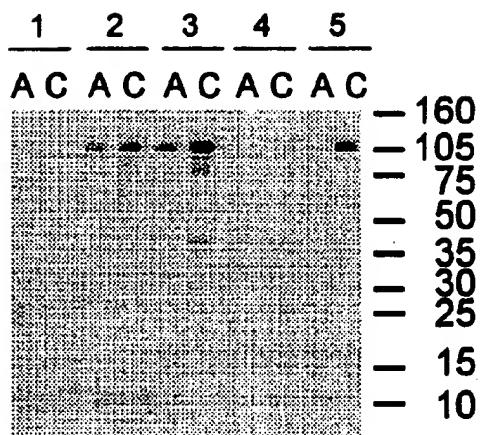
B. N-terminus
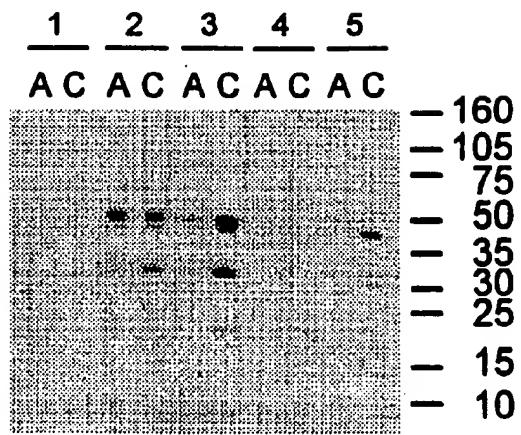
C. C-terminus
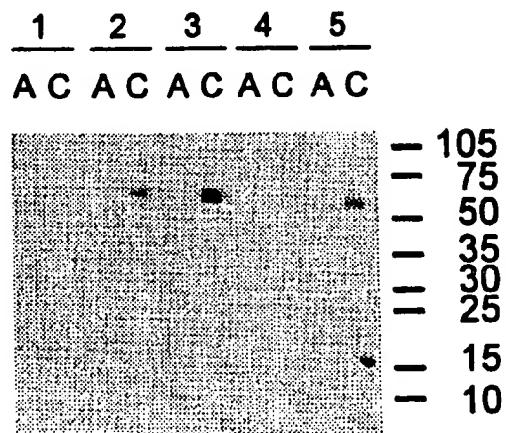
⛪ ExPASy Home page   **Site Map**   **Search ExPASy**   **Contact us**   **Swiss-Prot**

Search [Swiss-Prot/TrEMBL ▼] for [bvh-3 streptococcus] [Go] [Clear]

UniProtKB/TrEMBL entry Q9ANY1

[Printer-friendly view] [Request update] [Qt]

[Entry info] [Name and origin] [References] [Comments] [Cross-references] [Keywords] [Features] [Sequence] [Tools]

Note: most headings are clickable, even if they don't appear as links. They link to the user manual or other documents.

Entry information

Entry name	**Q9ANY1_STRPN**
Primary accession number	**Q9ANY1**
Secondary accession number	Q7D4B6
Entered in TrEMBL in	Release 17, June 2001
Sequence was last modified in	Release 17, June 2001
Annotations were last modified in	Release 30, May 2005

Name and origin of the protein

Protein name	**Pneumococcal histidine triad protein E [Precursor]**
Synonym	**Hypothetical protein SP1004**
Gene name	**Name: phtE** OrderedLocusNames: SP1004
From	Streptococcus pneumoniae [TaxID: 1313]
Taxonomy	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.

References

[1] NUCLEOTIDE SEQUENCE.
DOI=10.1128/IAI.69.2.949-958.2001; PubMed=11159990 [NCBI, ExPASy, EBI, Israel, Japan]
Adamou J.E., Heinrichs J.H., Erwin A.L., Walsh W., Gayle T., Dormitzer M., Dagan R., Brewah Y.A., Barren P., Lathigra R., Langermann S., Koenig S., Johnson S.;
"Identification and characterization of a novel family of pneumococcal proteins (the Pht family) that are protective against sepsis.";
Infect. Immun. 69:949-958(2001).

[2] NUCLEOTIDE SEQUENCE.
STRAIN=ATCC BAA-334 / TIGR4;
DOI=10.1126/science.1061217; PubMed=11463916 [NCBI, ExPASy, EBI, Israel, Japan]
Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D., Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J., Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D., Umayam L.A., White O., Salzberg S.L., Lewis M.R., ▓, Fraser C.M.;
"Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.";
Science 293:498-506(2001).

Comments

None

Cross-references

AF318956; AAK06761.1; -; Genomic_DNA.	[EMBL / GenBank / DDBJ] [CoDingSequence]

EMBL	AE007403; AAK75121.1; -; Genomic_DNA.	[EMBL / GenBank / DDBJ] [CoDingSequence]
PIR	H95115; H95115.	
TIGR	SP1004; -.	
InterPro	IPR006270; Strep_his_triad. Graphical view of domain structure.	
Pfam	PF04270; Strep_his_triad; 5. Pfam graphical view of domain structure.	
TIGRFAMs	TIGR01363; strep_his_triad; 3.	
ProDom	[Domain structure / List of seq. sharing at least 1 domain]	
HOGENOM	[Family / Alignment / Tree]	
ProtoMap	Q9ANY1.	
PRESAGE	Q9ANY1.	
ModBase	Q9ANY1.	
SWISS-2DPAGE	Get region on 2D PAGE.	
UniRef	View cluster of proteins with at least 50% / 90% identity.	

Keywords

Complete proteome; Hypothetical protein; Signal.

Features
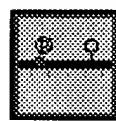
 Feature table viewer

Key	From	To	Length	Description
SIGNAL	1	29	29	Potential.

Sequence information

Length: **1039** AA   Molecular weight: **114631** Da   CRC64: **81A563FC806625C4** [This is a checksum on the sequence]

```
           10         20         30         40         50         60
    MKFSKKYIAA GSAVIVSLSL CAYALNQHRS QENKDNNRVS YVDGSQSSQK SENLTPDQVS

           70         80         90        100        110        120
    QKEGIQAEQI VIKITDQGYV TSHGDHYHYY NGKVPYDALF SEELLMKDPN YQLKDADIVN

          130        140        150        160        170        180
    EVKGGYIIKV DGKYYVYLKD AAHADNVRTK DEINRQKQEH VKDNEKVNSN VAVARSQGRY

          190        200        210        220        230        240
    TTNDGYVFNP ADIIEDTGNA YIVPHGGHYH YIPKSDLSAS ELAAAKAHLA GKNMQPSQLS

          250        260        270        280        290        300
    YSSTASDNNT QSVAKGSTSK PANKSENLQS LLKELYDSPS AQRYSESDGL VFDPAKIISR

          310        320        330        340        350        360
    TPNGVAIPHG DHYHFIPYSK LSALEEKIAR MVPISGTGST VSTNAKPNEV VSSLGSLSSN

          370        380        390        400        410        420
    PSSLTTSKEL SSASDGYIFN PKDIVEETAT AYIVRHGDHF HYIPKSNQIG QPTLPNNSLA

          430        440        450        460        470        480
```

```
      TPSPSLPINP GTSHEKHEED GYGFDANRII AEDESGFVMS HGDHNHYFFK KDLTEEQIKA

             490        500        510        520        530        540
      AQKHLEEVKT SHNGLDSLSS HEQDYPSNAK EMKDLDKKIE EKIAGIMKQY GVKRESIVVN

             550        560        570        580        590        600
      KEKNAIIYPH GDHHHADPID EHKPVGIGHS HSNYELFKPE EGVAKKEGNK VYTGEELTNV

             610        620        630        640        650        660
      VNLLKNSTFN NQNFTLANGQ KRVSFSFPPE LEKKLGINML VKLITPDGKV LEKVSGKVFG

             670        680        690        700        710        720
      EGVGNIANFE LDQPYLPGQT FKYTIASKDY PEVSYDGTFT VPTSLAYKMA SQTIFYPFHA

             730        740        750        760        770        780
      GDTYLRVNPQ FAVPKGTDAL VRVFDEFHGN AYLENNYKVG EIKLPIPKLN QGTTRTAGNK

             790        800        810        820        830        840
      IPVTFMANAY LDNQSTYIVE VPILEKENQT DKPSILPQFK RNKAQENLKL DEKVEEPKTS

             850        860        870        880        890        900
      EKVEKEKLSE TGNSTSNSTL EEVPTVDPVQ EKVAKFAESY GMKLENVLFN MDGTIELYLP

             910        920        930        940        950        960
      SGEVIKKNMA DFTGEAPQGN GENKPSENGK VSTGTVENQP TENKPADSLP EAPNEKPVKP

             970        980        990       1000       1010       1020
      ENSTDNGMLN PEGNVGSDPM LDPALEEAPA VDPVQEKLEK FTASYGLGLD SVIFNMDGTI

            1030
      ELRLPSGEVI KKNLSDLIA
```

Q9ANY1 in FASTA
format

View entry in original UniProtKB/TrEMBL format
View entry in raw text format (no links)
Request for annotation of this UniProtKB/TrEMBL entry

BLAST   BLAST submission on
         ExPASy/SIB
         or at NCBI (USA)

         ScanProsite, MotifScan

         NPSA Sequence analysis
         tools

Sequence analysis tools: ProtParam, ProtScale,
Compute pI/Mw, PeptideMass, PeptideCutter,
Dotlet (Java)

Submit a homology modeling request to SWISS-
MODEL

ExPASy Home page   **Site Map**   **Search ExPASy**   **Contact us**   Swiss-Prot
Hosted by ☰ NHRI Taiwan Mirror sites: Australia Brazil Canada Korea Switzerland USA

```
tr Q9ANY1        Pneumococcal histidine triad protein E precursor        1039
   Q9ANY1_STRPN  (Hypothetical                                           AA
                 protein SP1004) [phtE] [Streptococcus pneumoniae]       align
```

Score = 2006 bits (5196), Expect = 0.0
Identities = 998/1039 (96%), Positives = 1000/1039 (96%)

```
Query: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60
            MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS
Sbjct: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60

Query: 61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120
            QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN
Sbjct: 61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120

Query: 121  EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180
            EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY
Sbjct: 121  EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180

Query: 181  TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLS 240
            TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIP                     NMQPSQLS
Sbjct: 181  TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLS 240

Query: 241  YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300
            YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
Sbjct: 241  YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300

Query: 301  TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXX 360
            TPNGVAIPHGDHYHFIPYSKLSALEEKIAR VPISGTGSTVSTNAKPNEVV
Sbjct: 301  TPNGVAIPHGDHYHFIPYSKLSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSN 360

Query: 361  XXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420
                   KELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA
Sbjct: 361  PSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420

Query: 421  TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA 480
            TPSPSLPINPG SHEKHEEDGYGFDANRIIAEDESGF+MSHG+HNHYFFKKDLTEEQIKA
Sbjct: 421  TPSPSLPINPGTSHEKHEEDGYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKA 480

Query: 481  AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540
            AQKHLEEVKTSHNGLDSLSSHEQDYP NAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN
Sbjct: 481  AQKHLEEVKTSHNGLDSLSSHEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540

Query: 541  KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600
            KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
Sbjct: 541  KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600

Query: 601  VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660
            VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG
Sbjct: 601  VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660

Query: 661  EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720
            EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA
Sbjct: 661  EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720

Query: 721  GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780
            GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK
Sbjct: 721  GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780
```

```
Query: 781   IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS 840
             IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQEN KLDEKVEEPKTS
Sbjct: 781   IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENLKLDEKVEEPKTS 840

Query: 841   EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900
             EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
Sbjct: 841   EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900

Query: 901   SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960
             SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP
Sbjct: 901   SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960

Query: 961   ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020
             ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI
Sbjct: 961   ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020

Query: 1021  ELRLPSGEVIKKNLSDLIA 1039
             ELRLPSGEVIKKNLSDLIA
Sbjct: 1021  ELRLPSGEVIKKNLSDLIA 1039
```

```
tr  Q8DQ07        Pneumococcal histidine triad protein E [phtE]        1039
    Q8DQ07_STRR6  [Streptococcus                                        AA
                  pneumoniae (strain ATCC BAA-255 / R6)]               align
```

```
Score = 2017 bits (5225), Expect = 0.0
Identities = 1004/1039 (96%), Positives = 1004/1039 (96%)
```

```
Query: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60
            MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS
Sbjct: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60

Query: 61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120
            QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN
Sbjct: 61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120

Query: 121  EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180
            EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY
Sbjct: 121  EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180

Query: 181  TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQLS 240
            TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIP                      NMQPSQLS
Sbjct: 181  TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLS 240

Query: 241  YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300
            YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
Sbjct: 241  YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300

Query: 301  TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXX 360
            TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVV
Sbjct: 301  TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVSSLGSLSSN 360

Query: 361  XXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420
                   KELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA
Sbjct: 361  PSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420

Query: 421  TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA 480
            TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA
Sbjct: 421  TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA 480

Query: 481  AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540
            AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN
Sbjct: 481  AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540

Query: 541  KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600
            KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
Sbjct: 541  KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600

Query: 601  VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660
            VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG
Sbjct: 601  VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660

Query: 661  EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720
            EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA
Sbjct: 661  EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720

Query: 721  GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780
            GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK
Sbjct: 721  GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780
```

```
Query: 781   IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS 840
             IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS
Sbjct: 781   IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS 840

Query: 841   EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900
             EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
Sbjct: 841   EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900

Query: 901   SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960
             SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP
Sbjct: 901   SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960

Query: 961   ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020
             ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI
Sbjct: 961   ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020

Query: 1021  ELRLPSGEVIKKNLSDLIA 1039
             ELRLPSGEVIKKNLSDLIA
Sbjct: 1021  ELRLPSGEVIKKNLSDLIA 1039
```

CLUSTAL W (1.74) multiple sequence alignment


```
tr|Q9ANY1|Q9ANY1_STRPN      MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQK
tr|Q8DQ07|Q8DQ07_STRR6      MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQK
tr|Q6WNQ7|Q6WNQ7_STRPN      MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQK
tr|Q6WNQ5|Q6WNQ5_STRPN      --------------------CAYALNQHRSQENKDNNRVSYVDGSQSSQK
                                                *****************************


tr|Q9ANY1|Q9ANY1_STRPN      SENLTPDQVSQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALF
tr|Q8DQ07|Q8DQ07_STRR6      SENLTPDQVSQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALF
tr|Q6WNQ7|Q6WNQ7_STRPN      SENLTPDQVSQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALF
tr|Q6WNQ5|Q6WNQ5_STRPN      SENLTPDQVSQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALF
                            **************************************************


tr|Q9ANY1|Q9ANY1_STRPN      SEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADNVRTK
tr|Q8DQ07|Q8DQ07_STRR6      SEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADNVRTK
tr|Q6WNQ7|Q6WNQ7_STRPN      SEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADNVRTK
tr|Q6WNQ5|Q6WNQ5_STRPN      SEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADNVRTK
                            **************************************************


tr|Q9ANY1|Q9ANY1_STRPN      DEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA
tr|Q8DQ07|Q8DQ07_STRR6      DEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA
tr|Q6WNQ7|Q6WNQ7_STRPN      DEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA
tr|Q6WNQ5|Q6WNQ5_STRPN      DEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA
                            **************************************************


tr|Q9ANY1|Q9ANY1_STRPN      YIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSSTASDNNT
tr|Q8DQ07|Q8DQ07_STRR6      YIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSSTASDNNT
tr|Q6WNQ7|Q6WNQ7_STRPN      YIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSSTASDNNT
tr|Q6WNQ5|Q6WNQ5_STRPN      YIVPHRGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSSTASDNNT
                            ***** ********************************************


tr|Q9ANY1|Q9ANY1_STRPN      QSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
tr|Q8DQ07|Q8DQ07_STRR6      QSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
tr|Q6WNQ7|Q6WNQ7_STRPN      QSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
tr|Q6WNQ5|Q6WNQ5_STRPN      QSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
                            **************************************************


tr|Q9ANY1|Q9ANY1_STRPN      TPNGVAIPHGDHYHFIPYSKLSALEEKIARMVPISGTGSTVSTNAKPNEV
tr|Q8DQ07|Q8DQ07_STRR6      TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEV
tr|Q6WNQ7|Q6WNQ7_STRPN      TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEV
tr|Q6WNQ5|Q6WNQ5_STRPN      TPNGVAIPHGDHYHFIPYSKLSALEEKIARMVPISGTGSTVSTNAKPNEV
                            ****************************** *******************


tr|Q9ANY1|Q9ANY1_STRPN      VSSLGSLSSNPSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHF
tr|Q8DQ07|Q8DQ07_STRR6      VSSLGSLSSNPSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHF
tr|Q6WNQ7|Q6WNQ7_STRPN      VSSLGSLSSNPSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHF
tr|Q6WNQ5|Q6WNQ5_STRPN      VSSLGSLSSNPSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHF
                            **************************************************


tr|Q9ANY1|Q9ANY1_STRPN      HYIPKSNQIGQPTLPNNSLATPSPSLPINPGTSHEKHEEDGYGFDANRII
tr|Q8DQ07|Q8DQ07_STRR6      HYIPKSNQIGQPTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRII
tr|Q6WNQ7|Q6WNQ7_STRPN      HYIPKSNQIGQPTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRII
tr|Q6WNQ5|Q6WNQ5_STRPN      HYIPKSNQIGQPTLPNNSLATPSPSLPINPGTSHEKHEEDGYGFDANRII
                            ******************************* ******************


tr|Q9ANY1|Q9ANY1_STRPN      AEDESGFVMSHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS
tr|Q8DQ07|Q8DQ07_STRR6      AEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS
tr|Q6WNQ7|Q6WNQ7_STRPN      AEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS
```

```
tr|Q6WNQ5|Q6WNQ5_STRPN    AEDESGFVMSHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS
                          ******:****:**************************************

tr|Q9ANY1|Q9ANY1_STRPN    HEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPH
tr|Q8DQ07|Q8DQ07_STRR6    HEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPH
tr|Q6WNQ7|Q6WNQ7_STRPN    HEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPH
tr|Q6WNQ5|Q6WNQ5_STRPN    HEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPH
                          ******.*******************************************

tr|Q9ANY1|Q9ANY1_STRPN    GDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
tr|Q8DQ07|Q8DQ07_STRR6    GDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
tr|Q6WNQ7|Q6WNQ7_STRPN    GDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
tr|Q6WNQ5|Q6WNQ5_STRPN    GDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
                          **************************************************

tr|Q9ANY1|Q9ANY1_STRPN    VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKV
tr|Q8DQ07|Q8DQ07_STRR6    VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKV
tr|Q6WNQ7|Q6WNQ7_STRPN    VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKV
tr|Q6WNQ5|Q6WNQ5_STRPN    VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKV
                          **************************************************

tr|Q9ANY1|Q9ANY1_STRPN    LEKVSGKVFGEGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFT
tr|Q8DQ07|Q8DQ07_STRR6    LEKVSGKVFGEGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFT
tr|Q6WNQ7|Q6WNQ7_STRPN    LEKVSGKVFGEGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFT
tr|Q6WNQ5|Q6WNQ5_STRPN    LEKVSGKVFGEGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFT
                          **************************************************

tr|Q9ANY1|Q9ANY1_STRPN    VPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDALVRVFDEFHGN
tr|Q8DQ07|Q8DQ07_STRR6    VPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDALVRVFDEFHGN
tr|Q6WNQ7|Q6WNQ7_STRPN    VPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDALVRVFDEFHGN
tr|Q6WNQ5|Q6WNQ5_STRPN    VPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDALVRVFDEFHGN
                          **************************************************

tr|Q9ANY1|Q9ANY1_STRPN    AYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE
tr|Q8DQ07|Q8DQ07_STRR6    AYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE
tr|Q6WNQ7|Q6WNQ7_STRPN    AYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE
tr|Q6WNQ5|Q6WNQ5_STRPN    AYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE
                          **************************************************

tr|Q9ANY1|Q9ANY1_STRPN    VPILEKENQTDKPSILPQFKRNKAQENLKLDEKVEEPKTSEKVEKEKLSE
tr|Q8DQ07|Q8DQ07_STRR6    VPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSE
tr|Q6WNQ7|Q6WNQ7_STRPN    VPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSE
tr|Q6WNQ5|Q6WNQ5_STRPN    VPILEKENQTDKPSILPQFKRNKAQENSKFDEKVEEPKTSEKVEKEKLSE
                          ***************************.:*********************

tr|Q9ANY1|Q9ANY1_STRPN    TGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
tr|Q8DQ07|Q8DQ07_STRR6    TGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
tr|Q6WNQ7|Q6WNQ7_STRPN    TGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
tr|Q6WNQ5|Q6WNQ5_STRPN    TGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
                          **************************************************

tr|Q9ANY1|Q9ANY1_STRPN    SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLP
tr|Q8DQ07|Q8DQ07_STRR6    SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLP
tr|Q6WNQ7|Q6WNQ7_STRPN    SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLP
tr|Q6WNQ5|Q6WNQ5_STRPN    SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLP
                          **************************************************

tr|Q9ANY1|Q9ANY1_STRPN    EAPNEKPVKPENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEK
tr|Q8DQ07|Q8DQ07_STRR6    EAPNEKPVKPENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEK
tr|Q6WNQ7|Q6WNQ7_STRPN    EAPNEKPVKPENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEK
```

```
tr|Q6WNQ5|Q6WNQ5_STRPN        EAPNEKPVKPENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEK
                              *************************************************

tr|Q9ANY1|Q9ANY1_STRPN        FTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA
tr|Q8DQ07|Q8DQ07_STRR6        FTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA
tr|Q6WNQ7|Q6WNQ7_STRPN        FTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA
tr|Q6WNQ5|Q6WNQ5_STRPN        FTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA
                              **************************************
```

PileUp


```
   MSF: 1039  Type: P    Check:    304   ..

 Name: tr|Q9ANY1|Q9ANY1_STRPN oo  Len: 1039  Check: 9358  Weight: 0.100
 Name: tr|Q8DQ07|Q8DQ07_STRR6 oo  Len: 1039  Check: 8867  Weight: 0.100
 Name: tr|Q6WNQ7|Q6WNQ7_STRPN oo  Len: 1039  Check: 8867  Weight: 0.100
 Name: tr|Q6WNQ5|Q6WNQ5_STRPN oo  Len: 1039  Check: 3212  Weight: 0.100

//
```


```
tr|Q9ANY1|Q9ANY1_STRPN       MKFSKKYIAA GSAVIVSLSL CAYALNQHRS QENKDNNRVS YVDGSQSSQK
tr|Q8DQ07|Q8DQ07_STRR6       MKFSKKYIAA GSAVIVSLSL CAYALNQHRS QENKDNNRVS YVDGSQSSQK
tr|Q6WNQ7|Q6WNQ7_STRPN       MKFSKKYIAA GSAVIVSLSL CAYALNQHRS QENKDNNRVS YVDGSQSSQK
tr|Q6WNQ5|Q6WNQ5_STRPN       .......... .......... CAYALNQHRS QENKDNNRVS YVDGSQSSQK


tr|Q9ANY1|Q9ANY1_STRPN       SENLTPDQVS QKEGIQAEQI VIKITDQGYV TSHGDHYHYY NGKVPYDALF
tr|Q8DQ07|Q8DQ07_STRR6       SENLTPDQVS QKEGIQAEQI VIKITDQGYV TSHGDHYHYY NGKVPYDALF
tr|Q6WNQ7|Q6WNQ7_STRPN       SENLTPDQVS QKEGIQAEQI VIKITDQGYV TSHGDHYHYY NGKVPYDALF
tr|Q6WNQ5|Q6WNQ5_STRPN       SENLTPDQVS QKEGIQAEQI VIKITDQGYV TSHGDHYHYY NGKVPYDALF


tr|Q9ANY1|Q9ANY1_STRPN       SEELLMKDPN YQLKDADIVN EVKGGYIIKV DGKYYVYLKD AAHADNVRTK
tr|Q8DQ07|Q8DQ07_STRR6       SEELLMKDPN YQLKDADIVN EVKGGYIIKV DGKYYVYLKD AAHADNVRTK
tr|Q6WNQ7|Q6WNQ7_STRPN       SEELLMKDPN YQLKDADIVN EVKGGYIIKV DGKYYVYLKD AAHADNVRTK
tr|Q6WNQ5|Q6WNQ5_STRPN       SEELLMKDPN YQLKDADIVN EVKGGYIIKV DGKYYVYLKD AAHADNVRTK


tr|Q9ANY1|Q9ANY1_STRPN       DEINRQKQEH VKDNEKVNSN VAVARSQGRY TTNDGYVFNP ADIIEDTGNA
tr|Q8DQ07|Q8DQ07_STRR6       DEINRQKQEH VKDNEKVNSN VAVARSQGRY TTNDGYVFNP ADIIEDTGNA
tr|Q6WNQ7|Q6WNQ7_STRPN       DEINRQKQEH VKDNEKVNSN VAVARSQGRY TTNDGYVFNP ADIIEDTGNA
tr|Q6WNQ5|Q6WNQ5_STRPN       DEINRQKQEH VKDNEKVNSN VAVARSQGRY TTNDGYVFNP ADIIEDTGNA


tr|Q9ANY1|Q9ANY1_STRPN       YIVPHGGHYH YIPKSDLSAS ELAAAKAHLA GKNMQPSQLS YSSTASDNNT
tr|Q8DQ07|Q8DQ07_STRR6       YIVPHGGHYH YIPKSDLSAS ELAAAKAHLA GKNMQPSQLS YSSTASDNNT
tr|Q6WNQ7|Q6WNQ7_STRPN       YIVPHGGHYH YIPKSDLSAS ELAAAKAHLA GKNMQPSQLS YSSTASDNNT
tr|Q6WNQ5|Q6WNQ5_STRPN       YIVPHRGHYH YIPKSDLSAS ELAAAKAHLA GKNMQPSQLS YSSTASDNNT


tr|Q9ANY1|Q9ANY1_STRPN       QSVAKGSTSK PANKSENLQS LLKELYDSPS AQRYSESDGL VFDPAKIISR
tr|Q8DQ07|Q8DQ07_STRR6       QSVAKGSTSK PANKSENLQS LLKELYDSPS AQRYSESDGL VFDPAKIISR
tr|Q6WNQ7|Q6WNQ7_STRPN       QSVAKGSTSK PANKSENLQS LLKELYDSPS AQRYSESDGL VFDPAKIISR
tr|Q6WNQ5|Q6WNQ5_STRPN       QSVAKGSTSK PANKSENLQS LLKELYDSPS AQRYSESDGL VFDPAKIISR


tr|Q9ANY1|Q9ANY1_STRPN       TPNGVAIPHG DHYHFIPYSK LSALEEKIAR MVPISGTGST VSTNAKPNEV
tr|Q8DQ07|Q8DQ07_STRR6       TPNGVAIPHG DHYHFIPYSK LSALEEKIAR RVPISGTGST VSTNAKPNEV
tr|Q6WNQ7|Q6WNQ7_STRPN       TPNGVAIPHG DHYHFIPYSK LSALEEKIAR RVPISGTGST VSTNAKPNEV
tr|Q6WNQ5|Q6WNQ5_STRPN       TPNGVAIPHG DHYHFIPYSK LSALEEKIAR MVPISGTGST VSTNAKPNEV


tr|Q9ANY1|Q9ANY1_STRPN       VSSLGSLSSN PSSLTTSKEL SSASDGYIFN PKDIVEETAT AYIVRHGDHF
tr|Q8DQ07|Q8DQ07_STRR6       VSSLGSLSSN PSSLTTSKEL SSASDGYIFN PKDIVEETAT AYIVRHGDHF
tr|Q6WNQ7|Q6WNQ7_STRPN       VSSLGSLSSN PSSLTTSKEL SSASDGYIFN PKDIVEETAT AYIVRHGDHF
```

```
tr|Q6WNQ5|Q6WNQ5_STRPN        VSSLGSLSSN PSSLTTSKEL SSASDGYIFN PKDIVEETAT AYIVRHGDHF


tr|Q9ANY1|Q9ANY1_STRPN        HYIPKSNQIG QPTLPNNSLA TPSPSLPINP GTSHEKHEED GYGFDANRII
tr|Q8DQ07|Q8DQ07_STRR6        HYIPKSNQIG QPTLPNNSLA TPSPSLPINP GISHEKHEED GYGFDANRII
tr|Q6WNQ7|Q6WNQ7_STRPN        HYIPKSNQIG QPTLPNNSLA TPSPSLPINP GISHEKHEED GYGFDANRII
tr|Q6WNQ5|Q6WNQ5_STRPN        HYIPKSNQIG QPTLPNNSLA TPSPSLPINP GTSHEKHEED GYGFDANRII


tr|Q9ANY1|Q9ANY1_STRPN        AEDESGFVMS HGDHNHYFFK KDLTEEQIKA AQKHLEEVKT SHNGLDSLSS
tr|Q8DQ07|Q8DQ07_STRR6        AEDESGFIMS HGNHNHYFFK KDLTEEQIKA AQKHLEEVKT SHNGLDSLSS
tr|Q6WNQ7|Q6WNQ7_STRPN        AEDESGFIMS HGNHNHYFFK KDLTEEQIKA AQKHLEEVKT SHNGLDSLSS
tr|Q6WNQ5|Q6WNQ5_STRPN        AEDESGFVMS HGDHNHYFFK KDLTEEQIKA AQKHLEEVKT SHNGLDSLSS


tr|Q9ANY1|Q9ANY1_STRPN        HEQDYPSNAK EMKDLDKKIE EKIAGIMKQY GVKRESIVVN KEKNAIIYPH
tr|Q8DQ07|Q8DQ07_STRR6        HEQDYPGNAK EMKDLDKKIE EKIAGIMKQY GVKRESIVVN KEKNAIIYPH
tr|Q6WNQ7|Q6WNQ7_STRPN        HEQDYPGNAK EMKDLDKKIE EKIAGIMKQY GVKRESIVVN KEKNAIIYPH
tr|Q6WNQ5|Q6WNQ5_STRPN        HEQDYPSNAK EMKDLDKKIE EKIAGIMKQY GVKRESIVVN KEKNAIIYPH


tr|Q9ANY1|Q9ANY1_STRPN        GDHHHADPID EHKPVGIGHS HSNYELFKPE EGVAKKEGNK VYTGEELTNV
tr|Q8DQ07|Q8DQ07_STRR6        GDHHHADPID EHKPVGIGHS HSNYELFKPE EGVAKKEGNK VYTGEELTNV
tr|Q6WNQ7|Q6WNQ7_STRPN        GDHHHADPID EHKPVGIGHS HSNYELFKPE EGVAKKEGNK VYTGEELTNV
tr|Q6WNQ5|Q6WNQ5_STRPN        GDHHHADPID EHKPVGIGHS HSNYELFKPE EGVAKKEGNK VYTGEELTNV


tr|Q9ANY1|Q9ANY1_STRPN        VNLLKNSTFN NQNFTLANGQ KRVSFSFPPE LEKKLGINML VKLITPDGKV
tr|Q8DQ07|Q8DQ07_STRR6        VNLLKNSTFN NQNFTLANGQ KRVSFSFPPE LEKKLGINML VKLITPDGKV
tr|Q6WNQ7|Q6WNQ7_STRPN        VNLLKNSTFN NQNFTLANGQ KRVSFSFPPE LEKKLGINML VKLITPDGKV
tr|Q6WNQ5|Q6WNQ5_STRPN        VNLLKNSTFN NQNFTLANGQ KRVSFSFPPE LEKKLGINML VKLITPDGKV


tr|Q9ANY1|Q9ANY1_STRPN        LEKVSGKVFG EGVGNIANFE LDQPYLPGQT FKYTIASKDY PEVSYDGTFT
tr|Q8DQ07|Q8DQ07_STRR6        LEKVSGKVFG EGVGNIANFE LDQPYLPGQT FKYTIASKDY PEVSYDGTFT
tr|Q6WNQ7|Q6WNQ7_STRPN        LEKVSGKVFG EGVGNIANFE LDQPYLPGQT FKYTIASKDY PEVSYDGTFT
tr|Q6WNQ5|Q6WNQ5_STRPN        LEKVSGKVFG EGVGNIANFE LDQPYLPGQT FKYTIASKDY PEVSYDGTFT


tr|Q9ANY1|Q9ANY1_STRPN        VPTSLAYKMA SQTIFYPFHA GDTYLRVNPQ FAVPKGTDAL VRVFDEFHGN
tr|Q8DQ07|Q8DQ07_STRR6        VPTSLAYKMA SQTIFYPFHA GDTYLRVNPQ FAVPKGTDAL VRVFDEFHGN
tr|Q6WNQ7|Q6WNQ7_STRPN        VPTSLAYKMA SQTIFYPFHA GDTYLRVNPQ FAVPKGTDAL VRVFDEFHGN
tr|Q6WNQ5|Q6WNQ5_STRPN        VPTSLAYKMA SQTIFYPFHA GDTYLRVNPQ FAVPKGTDAL VRVFDEFHGN


tr|Q9ANY1|Q9ANY1_STRPN        AYLENNYKVG EIKLPIPKLN QGTTRTAGNK IPVTFMANAY LDNQSTYIVE
tr|Q8DQ07|Q8DQ07_STRR6        AYLENNYKVG EIKLPIPKLN QGTTRTAGNK IPVTFMANAY LDNQSTYIVE
tr|Q6WNQ7|Q6WNQ7_STRPN        AYLENNYKVG EIKLPIPKLN QGTTRTAGNK IPVTFMANAY LDNQSTYIVE
tr|Q6WNQ5|Q6WNQ5_STRPN        AYLENNYKVG EIKLPIPKLN QGTTRTAGNK IPVTFMANAY LDNQSTYIVE


tr|Q9ANY1|Q9ANY1_STRPN        VPILEKENQT DKPSILPQFK RNKAQENLKL DEKVEEPKTS EKVEKEKLSE
tr|Q8DQ07|Q8DQ07_STRR6        VPILEKENQT DKPSILPQFK RNKAQENSKL DEKVEEPKTS EKVEKEKLSE
tr|Q6WNQ7|Q6WNQ7_STRPN        VPILEKENQT DKPSILPQFK RNKAQENSKL DEKVEEPKTS EKVEKEKLSE
tr|Q6WNQ5|Q6WNQ5_STRPN        VPILEKENQT DKPSILPQFK RNKAQENSKF DEKVEEPKTS EKVEKEKLSE


tr|Q9ANY1|Q9ANY1_STRPN        TGNSTSNSTL EEVPTVDPVQ EKVAKFAESY GMKLENVLFN MDGTIELYLP
tr|Q8DQ07|Q8DQ07_STRR6        TGNSTSNSTL EEVPTVDPVQ EKVAKFAESY GMKLENVLFN MDGTIELYLP
tr|Q6WNQ7|Q6WNQ7_STRPN        TGNSTSNSTL EEVPTVDPVQ EKVAKFAESY GMKLENVLFN MDGTIELYLP
```

```
tr|Q6WNQ5|Q6WNQ5_STRPN      TGNSTSNSTL EEVPTVDPVQ EKVAKFAESY GMKLENVLFN MDGTIELYLP


tr|Q9ANY1|Q9ANY1_STRPN      SGEVIKKNMA DFTGEAPQGN GENKPSENGK VSTGTVENQP TENKPADSLP
tr|Q8DQ07|Q8DQ07_STRR6      SGEVIKKNMA DFTGEAPQGN GENKPSENGK VSTGTVENQP TENKPADSLP
tr|Q6WNQ7|Q6WNQ7_STRPN      SGEVIKKNMA DFTGEAPQGN GENKPSENGK VSTGTVENQP TENKPADSLP
tr|Q6WNQ5|Q6WNQ5_STRPN      SGEVIKKNMA DFTGEAPQGN GENKPSENGK VSTGTVENQP TENKPADSLP .


tr|Q9ANY1|Q9ANY1_STRPN      EAPNEKPVKP ENSTDNGMLN PEGNVGSDPM LDPALEEAPA VDPVQEKLEK
tr|Q8DQ07|Q8DQ07_STRR6      EAPNEKPVKP ENSTDNGMLN PEGNVGSDPM LDPALEEAPA VDPVQEKLEK
tr|Q6WNQ7|Q6WNQ7_STRPN      EAPNEKPVKP ENSTDNGMLN PEGNVGSDPM LDPALEEAPA VDPVQEKLEK
tr|Q6WNQ5|Q6WNQ5_STRPN      EAPNEKPVKP ENSTDNGMLN PEGNVGSDPM LDPALEEAPA VDPVQEKLEK


tr|Q9ANY1|Q9ANY1_STRPN      FTASYGLGLD SVIFNMDGTI ELRLPSGEVI KKNLSDLIA
tr|Q8DQ07|Q8DQ07_STRR6      FTASYGLGLD SVIFNMDGTI ELRLPSGEVI KKNLSDLIA
tr|Q6WNQ7|Q6WNQ7_STRPN      FTASYGLGLD SVIFNMDGTI ELRLPSGEVI KKNLSDLIA
tr|Q6WNQ5|Q6WNQ5_STRPN      FTASYGLGLD SVIFNMDGTI ELRLPSGEVI KKNLSDLIA
```